

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 8, 2002, 22:21:34 ; Search time 2186.81 Seconds
(without alignments)
92.580 Million cell updates/sec

Title: US-09-780-929-97

Perfect score: 15

Sequence: 1 agauaacyugaagu 15

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size: 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Listing first 45 summaries

Database:

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_hiv:*
15: em_gss_pin:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	264	9	AV214404 AV214404
2	15	100.0	270	9	BB373276 BB373276
3	15	100.0	398	9	AV673694 AV673694
4	15	100.0	410	12	AQ617558 AQ617558
5	15	100.0	565	9	AV861222 AV861222
6	15	100.0	572	9	AV890567 AV890567
7	15	100.0	589	9	AV883538 AV883538
8	15	100.0	614	9	AV840099 AV840099
9	15	100.0	615	9	AV844219 AV844219
10	15	100.0	640	9	AV889611 AV889611
11	15	100.0	678	9	AV681305 AV681305
12	15	100.0	682	9	AV866743 AV866743
13	15	100.0	688	9	AV842929 AV842929
14	15	100.0	699	10	BG441405 BG441405
15	15	100.0	729	9	AL663169 AL663169
16	15	100.0	729	9	AL666146 AL666146
17	15	100.0	795	10	BF784990 BF784990

c 18	15	100.0	824	9	BE052200 GA_Ea003
c 19	14	93.3	278	9	BB373413 BB373413
c 20	14	93.3	300	10	Z17985 ATT50469 AC
c 21	14	93.3	303	12	BH617107 SALK_0359
c 22	14	93.3	357	10	B1075246 IPL_17_H0
c 23	14	93.3	426	10	B1130325 G103P90Y
c 24	14	93.3	437	12	FR0002465
c 25	14	93.3	479	12	B26132
c 26	14	93.3	488	12	CNS00PCZ
c 27	14	93.3	493	12	TA344H01P
c 28	14	93.3	512	12	AZ926009 476.d1005
c 29	14	93.3	519	12	A0177762
c 30	14	93.3	522	12	A2629759
c 31	14	93.3	522	12	A2927699
c 32	14	93.3	534	12	A0971795
c 33	14	93.3	543	12	A0659631
c 34	14	93.3	546	10	B1131028
c 35	14	93.3	554	10	BM189295
c 36	14	93.3	559	12	TA45A11P
c 37	14	93.3	568	9	AA592673
c 38	14	93.3	596	9	A1981729
c 39	14	93.3	616	9	AM671648
c 40	14	93.3	619	12	A0639907
c 41	14	93.3	649	12	A0651706
c 42	14	93.3	727	10	B1258266
c 43	14	93.3	732	9	AL562830
c 44	14	93.3	749	12	A2825884
c 45	14	93.3	836	9	AL526032

ALIGNMENTS

RESULT 1	AV214404	264 bp	mrna	linear	EST 30-OCT-1999
LOCUS	AV214404				
DEFINITION	AV214404 RIKEN full-length enriched, ES cells Mus musculus cDNA				
ACCESSION	AV214404				
VERSION	AV214404.1	GI:6155250			
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Kono, H., Alzawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirokane, T., Horii, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, T., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tatenoe, M., Tomioka, Y., Tsunoda, Y., Watanabe, S., Yamamoto, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.				
COMMENT	RIKEN Mouse ESTs (Kono, H., et al. 1999)				
TITLE	Unpublished (1999)				
JOURNAL	Contact: Yoshihide Hayashizaki				
COMMENT	Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Saitama-cho, Tsukuba, Ibaraki, Japan Tel: 81-45-503-9212 Fax: 81-45-503-9216 Email: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ Sasaki, N., Izawa, M., Watanabe, M., Okazaki, Y., and Hayashizaki, Y. Matsuyama, S., Carninci, P., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y. Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998) Itoh, M., Katsunari, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,				

Tomaru, Y., Carinini, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashiaki, Y.
Automated filtration-based high-throughput plasmid preparation system. *Genome Res.* 9 (5), 463-470 (1999)
Carinini, P., and Hayashiaki, Y.
High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303, 19-44 (1999)

FEATURES

Source

BASE COUNTY
ORIGIN

```

Location/Qualifiers
1. 264
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_1b="Riken"
/clone_1lb="Riken full-length enriched, ES cells"
/cell_type="ES cells"
/lab_host="SOLR"
/note="Site_1: XhoI. Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGGAGAGAGAGATCCACAGCGCTCTTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 5.0 and subtraction to Rot = 25.0. Second strand
cDNA was prepared with the primer adapter of sequence [5'
GAGGAGAGAGATTCGAGGTAAATTAATATCCCCCCCCCCCCC 3']".

```

Query Match 100.0%; Score 15; DB 9; Length 264;
Best Local Similarity 80.0%; Pred. No. 40;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 aguaaacgugaagu 15
    |||:||||:||||:
Db 103 AGATAACGTGAAGT 117
```

RESULT	2
BB373276/c	
LOCUS	270 bp mRNA linear EST 13-JUL-2000
DEFINITION	BB373276 RIKEN full-length enriched, 16 days embryo head Mus musculus cDNA Cl30069K16 3', mRNA sequence.

ACCESSION	BB373276
VERSION	BB373276.1
	GI:9085770

KEYWORDS

5

SOURCE

ORGANISM M

三、四

REFERENCES

REFERENCE

TITLE	RIKEN Mouse ESTs (Konno, H., et al.)
JOURNAL	Unpublished (2000)
COMMENT	Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsic.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoke, S., Sasakawa, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Thermobisulfite and thermocatalyzed DNA methylation by trehalase and its application for the synthesis of full length cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 95 (2), 520-524 (1998)
Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y. Automated filtration-based high-throughput plasmid preparation system. *Genome Res.* 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303, 19-44 (1999)
Please visit our web site (<http://genome.etc.riken.go.jp>) for further details.

FEATURES

Source

```

Location/Qualifiers
1. .270
   /organism="Mus musculus"
   /strain="C57BL/6J"
   /db_xref="taxon:10090"
   /clone="Cl30069K16"
   /clone_lib="RIKEN full-length enriched, 16 days embryo
head"
   /sex="mixed"
   /tissue_type="head"
   /dev_stage="16 days embryo"
   /lab_host="DH10B"
/note="Site_1: Salt; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group In Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGGAGAGAGAGATCCAGAAGCTCTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
c-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGGAGAGAGATTCGAGTTRATTATTAATATCCCCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pluescript KS(+) after bulk excision from Lambda
FLC I"
```

BASE COUNT	71 a	88 c	42 g	69 t
ORIGIN				

ORIGIN

Query Match	100.0%	Score 15;	DB 9;	Length 270;
Best Local Similarity	80.0%	Pred. No. 40;		
Matches 12; Conservative	3;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	aga	uaa	cgu	ga	aga	au	15
Db	63	AGATAACGTGAAGAT	49					

RESULT	3				
AV673694					
LOCUS	AV673694	398 bp	mRNA	linear	EST-05-OCT-2000
DEFINITION	AV673694 Nori Satoh unpublished cDNA library c10a intestinalis				
	cDNA clone c10b1307 5', mRNA sequence.				

ACCESSION	AV673694
VERSION	AV673694.1
KEYWORDS	EST.
SOURCE	Ciona intestinalis.

ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.

REFERENCE 1 (bases 1 to 398)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished (2000)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
Source
1. 398
/organism="Ciona intestinalis"
/db_xref="taxon:7719"
/clone="c1b1307"
/clone_1lb="Nori Satoh unpublished cDNA library"
/tissue_type="whole animal"
/dev_stage="tailbud"

BASE COUNT 107 a 77 c 112 g 102 t

ORIGIN

Query Match 100.0%; Score 15; DB 9; Length 398;
Best Local Similarity 80.0%; Pred. No. 43;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaagugagaau 15
|||||:|||||:
DB 140 AGATACGTGAAGAT 154

RESULT 4
LOCUS AO617558/c 410 bp DNA linear GSS 15-JUN-1999
DEFINITION HS_5163_A2_H02_SPEE RPCI-11 Human Male BAC library Homo sapiens
genomic clone plate=739 COL=4 Row=O, DNA sequence.
ACCESSION AO617558
VERSION AO617558.1 GI:5078834
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 410)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.bufileo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.bufileo.edu/ordering_bac.htm)
or from Research Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 739 row: 0 column: 4
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 410.
Location/Qualifiers

source
1. 410
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="plate=739 COL=4 Row=O"
/clone_1lb="RPCI-11 Human Male BAC library"
/sex="male"
/note="Vector: pBAC3.6; Site.1: EcoRI; Site.2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI. Methylase. Size selected DNA was cloned into the
pBAC3.6 vector at EcoRI sites"

BASE COUNT 105 a 94 c 61 g 150 t

ORIGIN

Query Match 100.0%; Score 15; DB 12; Length 410;
Best Local Similarity 80.0%; Pred. No. 43;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaagugagaau 15
|||||:|||||:
DB 385 AGATACGTGAAGAT 371

RESULT 5
LOCUS AV861222/c 565 bp mRNA linear EST 08-NOV-2001
DEFINITION AV861222 Nori Satoh unpublished cDNA library, egg Ciona
intestinalis cDNA clone rc1eg34b01 3', mRNA sequence.
ACCESSION AV861222
VERSION AV861222.1 GI:16848746
KEYWORDS EST.
SOURCE Ciona intestinalis.
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.
REFERENCE 1 (bases 1 to 565)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished (2000)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
Source
1. 565
/organism="Ciona intestinalis"
/db_xref="taxon:7719"
/clone="rc1eg34b01"
/clone_1lb="Nori Satoh unpublished cDNA library, egg"
/tissue_type="whole animal"
/dev_stage="egg"

BASE COUNT 165 a 138 c 94 g 166 t 2 others

ORIGIN

Query Match 100.0%; Score 15; DB 9; Length 565;
Best Local Similarity 80.0%; Pred. No. 45;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaagugagaau 15
|||||:|||||:
DB 550 AGATACGTGAAGAT 536

RESULT 6
LOCUS AV890567/c 572 bp mRNA linear EST 09-NOV-2001
DEFINITION AV890567 Nori Satoh unpublished cDNA library, cleavage stage embryo
Ciona intestinalis cDNA clone rc1c128p23 3', mRNA sequence.

ACCESSION AV890567.1 GI:16879648
 VERSION EST.
 KEYWORDS Ciona intestinalis.
 SOURCE Ciona intestinalis.
 ORGANISM Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Clonidae; Ciona.
 REFERENCE 1 (bases 1 to 572)
 AUTHORS Satoh, N., Satou, Y., Kohara, Y., and Shin-1, T.
 TITLE Expressed genes in Ciona intestinalis
 JOURNAL Unpublished (2000)
 COMMENT Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoheascidian.zool.kyoto-u.ac.jp.
 Location/Qualifiers
 1. 572
 /organism="Ciona intestinalis"
 /db_xref="taxon:7719"
 /clone="rcic128p23"
 /clone_11b="Nori Satoh unpublished cDNA library, cleavage stage embryo"
 /tissue_type="whole animal"
 /dev_stage="cleavage stage embryo"
 BASE COUNT 165 a 143 c 95 g 163 t
 ORIGIN
 Query Match 100.0%; Score 15; DB 9; Length 572;
 Best Local Similarity 80.0%; Pred. No. 46;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 aguaacgugaagau 15
 |||:||||:||||:
 Db 524 AGATAACGTGAGAT 510
 RESULT 7
 AV883538 589 bp mRNA linear EST 08-NOV-2001
 LOCUS AV883538 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
 DEFINITION intestinalis cDNA clone rcitb39i04 3', mRNA sequence.
 ACCESSION AV883538
 VERSION AV883538.1 GI:16871062
 KEYWORDS EST.
 SOURCE Ciona intestinalis.
 ORGANISM Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Clonidae; Ciona.
 REFERENCE 1 (bases 1 to 589)
 AUTHORS Satoh, N., Satou, Y., Kohara, Y., and Shin-1, T.
 TITLE Expressed genes in Ciona intestinalis
 JOURNAL Unpublished (2000)
 COMMENT Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoheascidian.zool.kyoto-u.ac.jp.
 Location/Qualifiers
 1. 589
 /organism="Ciona intestinalis"
 /db_xref="taxon:7719"
 /clone="rcitb39i04"
 /clone_11b="Nori Satoh unpublished cDNA library, tailbud embryo"
 /tissue_type="whole animal"
 /dev_stage="tailbud embryo"
 BASE COUNT 167 a 142 c 102 g 178 t

ORIGIN
 Query Match 100.0%; Score 15; DB 9; Length 589;
 Best Local Similarity 80.0%; Pred. No. 46;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 aguaacgugaagau 15
 |||:||||:||||:
 Db 556 AGATAACGTGAGAT 542
 RESULT 8
 AV840099 614 bp mRNA linear EST 07-NOV-2001
 LOCUS AV840099 Nori Satoh unpublished cDNA library, cleavage stage embryo
 DEFINITION Ciona intestinalis cDNA clone rcic102d18, mRNA sequence.
 ACCESSION AV840099
 VERSION AV840099.1 GI:16784250
 KEYWORDS EST.
 SOURCE Ciona intestinalis.
 ORGANISM Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Clonidae; Ciona.
 REFERENCE 1 (bases 1 to 614)
 AUTHORS Satoh, N., Satou, Y., Kohara, Y., and Shin-1, T.
 TITLE Expressed genes in Ciona intestinalis
 JOURNAL Unpublished (2000)
 COMMENT Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoheascidian.zool.kyoto-u.ac.jp.
 Location/Qualifiers
 1. 614
 /organism="Ciona intestinalis"
 /db_xref="taxon:7719"
 /clone="rcic102d18"
 /clone_11b="Nori Satoh unpublished cDNA library, cleavage stage embryo"
 /tissue_type="whole animal"
 /dev_stage="cleavage stage embryo"
 BASE COUNT 177 a 147 c 112 g 177 t
 ORIGIN
 Query Match 100.0%; Score 15; DB 9; Length 614;
 Best Local Similarity 80.0%; Pred. No. 46;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 aguaacgugaagau 15
 |||:||||:||||:
 Db 546 AGATAACGTGAGAT 532
 RESULT 9
 AV844219 615 bp mRNA linear EST 08-NOV-2001
 LOCUS AV844219 Nori Satoh unpublished cDNA library, cleavage stage embryo
 DEFINITION Ciona intestinalis cDNA clone rcic105i11 3', mRNA sequence.
 ACCESSION AV844219
 VERSION AV844219.1 GI:16821441
 KEYWORDS EST.
 SOURCE Ciona intestinalis.
 ORGANISM Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Clonidae; Ciona.
 REFERENCE 1 (bases 1 to 615)
 AUTHORS Satoh, N., Satou, Y., Kohara, Y., and Shin-1, T.
 TITLE Expressed genes in Ciona intestinalis
 JOURNAL Unpublished (2000)

COMMENT

Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES

source

1. 615
/organism="Clona intestinalis"
/db_xref="taxon:7719"
/clone="rcic105111"
/clone_lib="Nori Satoh unpublished CDNA library, cleavage stage embryo"
/tissue_type="whole animal"
/dev_stage="cleavage stage embryo"
176 a 146 c 116 g 175 t 2 others

BASE COUNT

Query Match 100.0%; Score 15; DB 9; Length 615;
Best Local Similarity 80.0%; Pred. No. 46;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 aganaacgugaagu 15

DB 549 AGATACGTGAAGAT 535

RESULT 10

AV889611/c 640 bp mRNA linear EST 09-NOV-2001
LOCUS AV889611 Nori Satoh unpublished CDNA library, cleavage stage embryo
DEFINITION Clona intestinalis cDNA clone rcic148f17 3', mRNA sequence.

ACCESSION AV889611 GI:16878706
KEYWORDS EST.
SOURCE Clona intestinalis.
ORGANISM Clona intestinalis.

REFERENCE Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
1 (bases 1 to 640)
Satoh, N., Satou, Y., Kohara, Y. and Shin-I, T.

EXPOSED GENES IN Clona intestinalis
Unpublished (2000)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

COMMENT

FEATURES

source

1. 640
/organism="Clona intestinalis"
/db_xref="taxon:7719"
/clone="rcic148f17"
/clone_lib="Nori Satoh unpublished CDNA library, cleavage stage embryo"
/tissue_type="whole animal"
/dev_stage="cleavage stage embryo"
188 a 147 c 119 g 185 t 1 others

BASE COUNT

Query Match 100.0%; Score 15; DB 9; Length 640;
Best Local Similarity 80.0%; Pred. No. 46;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 aganaacgugaagu 15

DB 569 AGATACGTGAAGAT 555

RESULT 11

AV681305 678 bp mRNA linear EST 05-OCT-2000
LOCUS AV681305 Nori Satoh unpublished CDNA library Clona intestinalis
DEFINITION cDNA clone rcic1b1925 3', mRNA sequence.

ACCESSION AV681305 GI:10119304
VERSION AV681305.1
KEYWORDS EST.
SOURCE Clona intestinalis.
ORGANISM Clona intestinalis.

REFERENCE Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
1 (bases 1 to 678)
Satoh, N., Satou, Y., Kohara, Y. and Shin-I, T.

EXPOSED GENES IN Clona intestinalis
Unpublished (2000)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

COMMENT

FEATURES

source

1. 678
/organism="Clona intestinalis"
/db_xref="taxon:7719"
/clone="rcic1b1925"
/clone_lib="Nori Satoh unpublished CDNA library"
/tissue_type="whole animal"
/dev_stage="tailbud"
200 a 120 c 153 g 205 t

BASE COUNT

Query Match 100.0%; Score 15; DB 9; Length 678;
Best Local Similarity 80.0%; Pred. No. 47;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 aganaacgugaagu 15

DB 88 AGATACGTGAAGAT 102

RESULT 12

AV866743/c 682 bp mRNA linear EST 08-NOV-2001
LOCUS AV866743 Nori Satoh unpublished CDNA library, egg Clona
DEFINITION intestinalis cDNA clone rcleg47106 3', mRNA sequence.

ACCESSION AV866743 GI:16854267
VERSION AV866743.1
KEYWORDS EST.
SOURCE Clona intestinalis.
ORGANISM Clona intestinalis.

REFERENCE Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
1 (bases 1 to 682)
Satoh, N., Satou, Y., Kohara, Y. and Shin-I, T.

EXPOSED GENES IN Clona intestinalis
Unpublished (2000)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

COMMENT

FEATURES

source

1. 682
/organism="Clona intestinalis"
/db_xref="taxon:7719"
/clone="rcleg47106"

```
/clone_lib="Nori Satoh unpublished cDNA library, egg"
/tissue_type="whole animal"
/dev_stage="egg"
BASE COUNT      205 a      158 c      117 g      199 t      3 others
ORIGIN

Query Match      100.0%; Score 15; DB 9; Length 682;
Best Local Similarity 80.0%; Pred. No. 47;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 agauaacgugaagu 15
      |||:||||:||||:
      Db      551 AGATTAACGTGAAGAT 537

RESULT 13
AV842929      688 bp      mRNA      linear      EST 08-NOV-2001
LOCUS      AV842929 Nori Satoh unpublished cDNA library, larva Clona
DEFINITION      intestinalis cDNA clone rcllv04p12 3', mRNA sequence.
ACCESSION      AV842929
VERSION      AV842929.1 GI:16818879
KEYWORDS      EST.
SOURCE      Clona intestinalis.
ORGANISM      Clona intestinalis.
Eukaryota; Metazoa; Chordata; Trochozoa; Ascidacea; Enterogona;
Phlebobranchia; Clonidae; Clona.
REFERENCE      1 (bases 1 to 688)
AUTHORS      Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
TITLE      Expressed genes in Clona intestinalis
JOURNAL      Unpublished (2000)
COMMENT      Contact: Nori Satoh
      Department of Zoology
      Kyoto University
      Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
      Tel: 81-75-753-4081
      Fax: 81-75-705-1113
      Email: satoh@ascidian.zool.kyoto-u.ac.jp.
      Location/Qualifiers
FEATURES
      source      1..688
      /organism="Clona intestinalis"
      /db_xref="taxon:7719"
      /clone="rcllv04p12"
      /clone_lib="Nori Satoh unpublished cDNA library, larva"
      /tissue_type="whole animal"
      /dev_stage="larva"
BASE COUNT      210 a      155 c      120 g      202 t      1 others
ORIGIN

Query Match      100.0%; Score 15; DB 9; Length 688;
Best Local Similarity 80.0%; Pred. No. 47;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 agauaacgugaagu 15
      |||:||||:||||:
      Db      568 AGATTAACGTGAAGAT 554

RESULT 14
BG441405      699 bp      mRNA      linear      EST 15-MAR-2001
LOCUS      BG441405 Ga_Ea0013D08f Gossypium arboreum 7-10 dpa fiber library Gossypium
DEFINITION      Ga_Ea0013D08f Gossypium arboreum 7-10 dpa fiber library Gossypium
ACCESSION      BG441405
VERSION      BG441405.1 GI:13351057
KEYWORDS      EST.
SOURCE      Gossypium arboreum.
ORGANISM      Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.

/clone_lib="Nori Satoh unpublished cDNA library, egg"
/tissue_type="whole animal"
/dev_stage="egg"
BASE COUNT      205 a      158 c      117 g      199 t      3 others
ORIGIN

Query Match      100.0%; Score 15; DB 10; Length 699;
Best Local Similarity 80.0%; Pred. No. 47;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 agauaacgugaagu 15
      |||:||||:||||:
      Db      669 AGATTAACGTGAAGAT 655

RESULT 15
AL663169      729 bp      mRNA      linear      EST 11-JAN-2002
LOCUS      AL663169 directional larval cDNA library Clona intestinalis cDNA
DEFINITION      Clona 002ZG09 3', mRNA sequence.
ACCESSION      AL663169
VERSION      AL663169.1 GI:18130077
KEYWORDS      EST.
SOURCE      Clona intestinalis.
ORGANISM      Clona intestinalis.
Eukaryota; Metazoa; Chordata; Trochozoa; Ascidacea; Enterogona;
Phlebobranchia; Clonidae; Clona.
REFERENCE      1 (bases 1 to 729)
AUTHORS      Unpublished (2002)
TITLE      Clona intestinalis directional larval cDNA library
JOURNAL      Unpublished (2002)
COMMENT      Contact: Genoscope
      Genoscope - Centre National de Sequencage
      BP 191 91006 EVRY cedex - France
      Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
      IMPORTANT: this sequence may contain errors. The Clona intestinalis
      library from which the clone was isolated may be contaminated with
      cDNAs from bacteria or other Eukarya.
      Directional larval cDNA library originate from Dr. M. Branno,
      Stazione A. Dohrn, Naples, Italy, and was prepared in
      pluescript2SK+.
FEATURES
      source      1..729
      /organism="Clona intestinalis"
      /db_xref="taxon:7719"
      /clone="002ZG09"
      /clone_lib="directional larval cDNA library"
      /note="Vector: pluescript2SK+"
REFERENCE      1 (bases 1 to 699)
AUTHORS      Wang,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
      ,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE      An integrated analysis of the genetics, development, and evolution
      of the cotton fiber
JOURNAL      Unpublished (2000)
COMMENT      Contact: Wang RA
      Clemson University Genomics Institute
      Clemson University
      100 Jordan Hall, Clemson, SC 29634, USA
      Tel: 864 656 7288
      Fax: 864 656 4293
      Email: rwing@clemson.edu
      Seq primer: TAATACGACTCACTATAGG
      High quality sequence stop: 694.
FEATURES
      source      1..699
      /organism="Gossypium arboreum"
      /strain="AKA"
      /cultivar="8400"
      /db_xref="taxon:29729"
      /clone="GA_Ea0013D08f"
      /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
      /tissue_type="Fibers isolated from bolls harvested 7-10
      dpa"
      /lab_host="E. coli"
      /note="Vector: pBK-CMV, Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      198 a      165 c      154 g      182 t
ORIGIN
```

BASE COUNT 223 a 159 c 126 g 217 t 4 others
ORIGIN

Query Match

100.0%; Score 15; DB 9; Length 729;

Best Local Similarity 80.0%; Pred. No. 47;

Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaagau 15

|||||:|||||:

Db 600 AGATAACGTGAAGAT 586

Search completed: September 9, 2002, 01:12:49
Job time: 10275 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 8, 2002, 22:22:35 : Search time 57.83 Seconds
(without alignments)
63.713 Million cell updates/sec

Title: US-09-780-929-97

Sequence: 15
1 agauacgugaau 15

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 38353 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cg2_6/prodata/2/lna/5A_COMB.seq:*
- 2: /cg2_6/prodata/2/lna/5B_COMB.seq:*
- 3: /cg2_6/prodata/2/lna/5A_COMB.seq:*
- 4: /cg2_6/prodata/2/lna/5B_COMB.seq:*
- 5: /cg2_6/prodata/2/lna/PCrUS_COMB.seq:*
- 6: /cg2_6/prodata/2/lna/ackfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	14	93.3	9620	4	US-08-952-127-11
C 2	12	80.0	231	1	US-08-450-834-1
C 3	12	80.0	504	3	US-08-765-332-113
C 4	12	80.0	504	4	US-09-448-894-113
C 5	12	80.0	701	4	US-08-998-416-830
C 6	12	80.0	762	4	US-08-998-416-462
C 7	12	80.0	831	1	US-08-450-834-5
C 8	12	80.0	1446	1	US-08-173-436A-1
C 9	12	80.0	1446	1	US-08-173-436A-3
C 10	12	80.0	2238	1	US-08-173-436A-4
C 11	12	80.0	4138	1	US-08-323-474-1
C 12	12	80.0	4138	5	PCT-US93-06093-1
C 13	12	80.0	5622	4	US-09-067-800-3
C 14	12	80.0	5622	4	US-09-349-677-3
C 15	12	80.0	9319	4	US-08-976-259-85
C 16	12	80.0	18318	1	US-08-414-926A-6
C 17	12	80.0	18318	2	US-08-926-922-6
C 18	12	80.0	18318	3	US-09-253-682-6
C 19	12	80.0	18318	4	US-09-527-657-6
C 20	12	80.0	22671	4	US-08-976-259-14
C 21	12	80.0	43676	3	US-09-356-952-12
C 22	11	73.3	144	2	US-08-469-537A-61
C 23	11	73.3	241	3	US-08-621-018B-16
C 24	11	73.3	242	3	US-08-621-018B-30
C 25	11	73.3	247	3	US-08-621-018B-23
C 26	11	73.3	251	3	US-08-621-018B-19
C 27	11	73.3	277	3	US-08-621-018B-27

28	11	73.3	284	3	US-08-621-018B-20	Sequence 20, Appl
29	11	73.3	309	1	US-08-409-373B-1	Sequence 1, Appl1
30	11	73.3	309	1	US-08-409-373B-1	Sequence 1, Appl1
31	11	73.3	309	3	US-08-621-018B-1	Sequence 114, Appl
C 32	11	73.3	499	3	US-08-765-332-114	Sequence 114, Appl
C 33	11	73.3	499	4	US-09-448-894-114	Sequence 28, Appl
C 34	11	73.3	562	4	US-08-976-259-28	Sequence 520, Appl
C 35	11	73.3	613	4	US-09-385-882-520	Sequence 811, Appl
C 36	11	73.3	710	4	US-08-469-537A-79	Sequence 151, Appl
37	11	73.3	755	2	US-08-961-083-151	Sequence 5, Appl
38	11	73.3	794	3	US-08-953-040-5	Sequence 227, Appl
39	11	73.3	815	4	US-09-439-313-227	Sequence 29, Appl
40	11	73.3	818	3	US-09-248-335-29	Sequence 25, Appl
41	11	73.3	1134	3	US-08-188-582-25	Sequence 7, Appl1
42	11	73.3	1183	1	US-08-646-715-25	
43	11	73.3	1314	4	US-08-800-682-1	
44	11	73.3	1341	2	US-08-945-848-7	
C 45	11	73.3	1341	2	US-08-945-848-7	

ALIGNMENTS

RESULT 1
US-08-952-127-11/c
Sequence 11, Application US/08952127
Patent No. 6211336
GENERAL INFORMATION:
APPLICANT: Shiloh, Yosef
APPLICANT: Tagle, Danilo A.
APPLICANT: Collins, Francis S.
TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
STREET: 30500 No. 6211336Western Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: U.S.
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/952;127
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,995
REFERENCE/DOCKET NUMBER: 2290.00029
TELECOMMUNICATION INFORMATION:
TELEPHONE: 810-539-5050
TELEFAX: 810-539-5050
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 9620 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Mus musculus
POSITION IN GENOME:
CHROMOSOME/SEGMENT: Chromosome 9, Band 9C
US-08-952-127-11

Query Match 93.3%; Score 14; DB 4; Length 9620;
Best Local Similarity 78.6%; Pred. No. 2.5;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 gaaacgugaag 15
|:||||:|
DB 4622 GATAAGTGAAG 4609

RESULT 2

US-08-450-834-1
; Sequence 1, Application US/08450834
; Patent No. 5773705
; GENERAL INFORMATION:
; APPLICANT: Vierstra, Richard D
; APPLICANT: Hondred, David
; APPLICANT: Callis, Judy
; TITLE OF INVENTION: Ubiquitin Fusion Protein System for
; TITLE OF INVENTION: Protein Production in Plants
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: P.O. Box 2113
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,834
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/999,709
; FILING DATE: 31-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 960296, 92425
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 231 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: C-terminal
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; IMMEDIATE SOURCE:
; CLONE: UBQ11
; US-08-450-834-1

Query Match 80.0%; Score 12; DB 1; Length 231;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 gaaacgugaag 13
|:||||:|
DB 70 GATAAGTGAAG 81

RESULT 3
US-08-765-332-113/C
; Sequence 113, Application US/08765332
; Patent No. 6025132

; GENERAL INFORMATION:
; APPLICANT: JANNES, GEERT
; APPLICANT: ROSSAU, RUDI
; APPLICANT: VAN HEUVERSWYN, HUGO
; TITLE OF INVENTION: SIMULTANEOUS DETECTION, IDENTIFICATION
; TITLE OF INVENTION: AND DIFFERENTIATION OF EUBACTERIAL TAXA USING A
; NUMBER OF SEQUENCES: 216
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,332
; FILING DATE: 23-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP95/02452
; FILING DATE: 23-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95870032.0
; FILING DATE: 07-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 94870106.5
; FILING DATE: 24-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4091
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-765-332-113

Query Match 80.0%; Score 12; DB 3; Length 504;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 auaacgugaaga 14
|:||||:|
DB 474 ATAAGTGAAGA 463

RESULT 4
US-09-448-894-113/C
; Sequence 113, Application US/09448894
; Patent No. 6312903
; GENERAL INFORMATION:
; APPLICANT: JANNES, GEERT
; APPLICANT: ROSSAU, RUDI
; APPLICANT: VAN HEUVERSWYN, HUGO
; TITLE OF INVENTION: SIMULTANEOUS DETECTION, IDENTIFICATION
; AND DIFFERENTIATION OF EUBACTERIAL TAXA USING A
; HYBRIDIZATION ASSAY
; NUMBER OF SEQUENCES: 216

CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P C
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/448,894
FILING DATE: 29-NOV-1993
CLASSIFICATION: <Unknown>
07-APR-1995
24-JUN-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/765,332
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 95870032.0
FILING DATE: 07-APR-1995
APPLICATION NUMBER: EP 94870106.5
FILING DATE: 24-JUN-1994

ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-14
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4091
TELEFAX: 703-816-4100

INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 113:
US-09-448-894-113

Query Match 80.0%; Score 12; DB 4; Length 504;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 auaacugaga 14
|:||||:||||
DB 474 ATACGTGAGA 463

RESULT 5
US-08-998-416-830/c
Sequence 830, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Redischung, Corine
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBVA GOSSTPIT
TITLE OF INVENTION: AND USUS THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park

STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 830:
SEQUENCE CHARACTERISTICS:
LENGTH: 701 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE: PAG1526UP
US-08-998-416-830

Query Match 80.0%; Score 12; DB 4; Length 701;
Best Local Similarity 75.0%; Pred. No. 40;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 uaacugaga 15
|:||||:||||
DB 73 TACGTGAGAT 62

RESULT 6
US-08-998-416-462
Sequence 462, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Redischung, Corine
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBVA GOSSTPIT
TITLE OF INVENTION: AND USUS THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997

```

? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: CH 0016/97
? FILING DATE: 31-DEC-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Meigs, J. Timothy
? REGISTRATION NUMBER: 38,241
? REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 919-541-8587
? TELEFAX: 919-541-8689
? INFORMATION FOR SEQ ID NO: 462:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 762 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? ORIGINAL SOURCE:
? ORGANISM: PAG1332RP
?
US-08-998-416-462

Query Match      80.0%; Score 12; DB 4; Length 762;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 aguaacgugaa 12
DB 210 AGATAACGTGAA 221

RESULT 7
US-08-450-834-5
? Sequence 5, Application US/08450834
? Patent No. 5773705
? GENERAL INFORMATION:
? APPLICANT: Viorstra, Richard D
? APPLICANT: Honored, David
? APPLICANT: Callis, Judy
? TITLE OF INVENTION: Ubiquitin Fusion Protein System for
? TITLE OF INVENTION: Protein Production in Plants
? NUMBER OF SEQUENCES: 6
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Quarles & Brady
? STREET: P.O. Box 2113
? CITY: Madison
? STATE: WI
? COUNTRY: USA
? ZIP: 53701-2113
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/450,834
? FILING DATE: 25-MAY-1995
? CLASSIFICATION: 800
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/999,709
? FILING DATE: 31-DEC-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: Seay, Nicholas J
? REGISTRATION NUMBER: 27,386
? REFERENCE/DOCKET NUMBER: 960296.92425
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 608-251-5000
? TELEFAX: 608-251-9166
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 831 base pairs
? TYPE: nucleic acid
```

```

? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: other nucleic acid
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? IMMEDIATE SOURCE:
? CLONE: 35S/AMV/UBQ11/UBQ-GUS
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 503..730
? FEATURE:
? NAME/KEY: promoter
? LOCATION: 1..502
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 1..6
? OTHER INFORMATION: /function="pco RI restriction
? OTHER INFORMATION: site"
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 7..12
? OTHER INFORMATION: /function="Sac I restriction site"
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 13..18
? OTHER INFORMATION: /function="Kpn I restriction site"
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 18..24
? OTHER INFORMATION: /function="Apa I restriction site"
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 464..469
? OTHER INFORMATION: /function="Hind III restriction
? OTHER INFORMATION: site"
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 720..726
? OTHER INFORMATION: /function="Sac II restriction
? OTHER INFORMATION: site"
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 819..825
? OTHER INFORMATION: /function="Bcl I restriction site"
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 826..831
? OTHER INFORMATION: /function="Xba I restriction site"
?
US-08-450-834-5

Query Match      80.0%; Score 12; DB 1; Length 831;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 gaaacgugaa 13
DB 572 GATAACGTGAAG 583

RESULT 8
US-08-173-436A-1/C
? Sequence 1, Application US/08173436A
? Patent No. 5698444
? GENERAL INFORMATION:
? APPLICANT: Baer, Melvyn
? APPLICANT: Kuisar, Jonathan D.
? TITLE OF INVENTION: SEROTONIN RECEPTOR PROTEIN AND RELATED
? TITLE OF INVENTION: NUCLEIC ACID COMPOUNDS
? NUMBER OF SEQUENCES: 4
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Eli Lilly and Company
? STREET: Lilly Corporate Center/Patent Division
```

CITY: Indianapolis
STATE: IN
COUNTRY: US
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/173,436A
FILING DATE: 23-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gaylo, Paul J.
REGISTRATION NUMBER: 36,808
REFERENCE/DOCKET NUMBER: X-9367
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-0756
TELEFAX: 317-276-3861
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1446 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1443
US-08-173-436A-1

Query Match 80.0%; Score 12; DB 1; Length 1446;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacugaa 12
|||:||||:|
DB 82 AGATAACGTGA 71

RESULT 9
US-08-173-436A-3/C
Sequence 3, Application US/08173436A
Patent No. 5698444
GENERAL INFORMATION:
APPLICANT: Baez, Melvyn
APPLICANT: Kursar, Jonathan D.
TITLE OF INVENTION: SEROTONIN RECEPTOR PROTEIN AND RELATED
TITLE OF INVENTION: NUCLEIC ACID COMPOUNDS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center/Patent Division
CITY: Indianapolis
STATE: IN
COUNTRY: US
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/173,436A
FILING DATE: 23-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gaylo, Paul J.
REGISTRATION NUMBER: 36,808
REFERENCE/DOCKET NUMBER: X-9367
TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-0756
TELEFAX: 317-276-3861
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1446 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: mRNA
US-08-173-436A-3

Query Match 80.0%; Score 12; DB 1; Length 1446;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacugaa 12
|||:||||:|
DB 82 AGATAACGTGA 71

RESULT 10
US-08-173-436A-4/C
Sequence 4, Application US/08173436A
Patent No. 5698444
GENERAL INFORMATION:
APPLICANT: Baez, Melvyn
APPLICANT: Kursar, Jonathan D.
TITLE OF INVENTION: SEROTONIN RECEPTOR PROTEIN AND RELATED
TITLE OF INVENTION: NUCLEIC ACID COMPOUNDS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center/Patent Division
CITY: Indianapolis
STATE: IN
COUNTRY: US
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/173,436A
FILING DATE: 23-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gaylo, Paul J.
REGISTRATION NUMBER: 36,808
REFERENCE/DOCKET NUMBER: X-9367
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-0756
TELEFAX: 317-276-3861
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2238 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-173-436A-4

Query Match 80.0%; Score 12; DB 1; Length 2238;
Best Local Similarity 83.3%; Pred. No. 41;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacugaa 12
|||:||||:|
DB 549 AGATAACGTGA 538

RESULT 11
US-08-323-474-1
; Sequence 1, Application US/08323474
; Patent No. 5447860
; GENERAL INFORMATION:
; APPLICANT: Ziegler, Steven F.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/323,474
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/905,600
; FILING DATE: 26-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2609
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4138 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 149..3523
; US-08-323-474-1

Query Match 80.0%; Score 12; DB 1; Length 4138;
Best Local Similarity 83.3%; Pred. No. 41;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 agauaacugaa 12
|||:||||:||||
Db 556 AGATACGTGAA 567

RESULT 12
PCT-US93-06093-1
; Sequence 1, Application PC/TUS9306093
; GENERAL INFORMATION:
; APPLICANT: Ziegler, Steven F.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06093
; FILING DATE: 19930625
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/905,600
; FILING DATE: 26-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2609
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4138 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 149..3523
; PCT-US93-06093-1

Query Match 80.0%; Score 12; DB 5; Length 4138;
Best Local Similarity 83.3%; Pred. No. 41;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 agauaacugaa 12
|||:||||:||||
Db 556 AGATACGTGAA 567

RESULT 13
US-09-067-800-3
; Sequence 3, Application US/09067800
; Patent No. 6198024
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; TITLE OF INVENTION: Seed Plants Characterized by Delayed
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/067,800
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Kathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 2948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5622 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..5622
OTHER INFORMATION: /label= AG11_promoter
US-09-067-800-3

Query Match
Best Local Similarity 80.0%; Score 12; DB 4; Length 5622;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaa 12
|||||:||||
DB 4984 AGATAAGTGAA 4995

RESULT 14
US-09-349-677-3
Sequence 3, Application US/09349677
Patent No. 6288305
GENERAL INFORMATION:
APPLICANT: Yanofsky, Martin F.
APPLICANT: Ferandiz, Cristina
TITLE OF INVENTION: Seed Plants Characterized by Delayed
TITLE OF INVENTION: Seed Dispersal
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/349,677
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/067,800
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UD 2948
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5622 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..5622
OTHER INFORMATION: /label= AG11_promoter
US-09-349-677-3

Query Match
Best Local Similarity 80.0%; Score 12; DB 4; Length 5622;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaa 12
|||||:||||
DB 4984 AGATAAGTGAA 4995

RESULT 15
US-08-976-259-85
Sequence 85, Application US/08976259
Patent No. 6316609
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Choi, Gil H.
APPLICANT: Welch, Rodney A.
TITLE OF INVENTION: Nucleotide Sequence of Escherichia coli
Patent No. 6316609
NUMBER OF SEQUENCES: 142
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Ave, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976,259
FILING DATE: Herewith
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,626 AND US 60/061,953
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0740002/EKS/CBM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 85:
SEQUENCE CHARACTERISTICS:
LENGTH: 9319 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-976-259-85

Query Match
Best Local Similarity 80.0%; Score 12; DB 4; Length 9319;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaa 12
|||||:||||
DB 4882 AGATAAGTGAA 4893

Search completed: September 9, 2002, 01:14:06
Job time: 10291 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 9, 2002, 00:36:00 ; Search time 268.97 Seconds
(without alignments)
95.749 Million cell updates/sec

Title: US-09-780-929-97

Perfect score: 15
Sequence: 1 agaaacgugaagau 15

Scoring table: OLIGO_MTC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N.Geneseq_032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	15	AA512347	DNA encoding deoxy
2	15	100.0	27	AA512295	DNA encoding class
3	15	100.0	27	AA512296	DNA encoding class
4	15	100.0	27	AA512287	DNA encoding class
5	15	100.0	27	AA512298	DNA encoding class
6	15	100.0	27	AA512299	DNA encoding class
7	15	100.0	27	AA512300	DNA encoding class
8	15	100.0	27	AA512301	DNA encoding class
9	15	100.0	27	AA512302	DNA encoding class

10	15	100.0	27	AA512303	DNA encoding class
11	15	100.0	27	AA512304	DNA encoding class
12	15	100.0	27	AA512305	DNA encoding class
13	15	100.0	27	AA512306	DNA encoding class
14	15	100.0	27	AA512307	DNA encoding class
15	15	100.0	27	AA512308	DNA encoding class
16	15	100.0	27	AA512309	DNA encoding class
17	15	100.0	27	AA512310	DNA encoding class
18	15	100.0	27	AA512311	DNA encoding class
19	15	100.0	27	AA512312	DNA encoding class
20	15	100.0	27	AA512313	DNA encoding class
21	15	100.0	27	AA512314	DNA encoding class
22	15	100.0	27	AA512315	DNA encoding class
23	15	100.0	27	AA512316	DNA encoding class
24	15	100.0	27	AA512317	DNA encoding class
25	15	100.0	27	AA512318	DNA encoding class
26	15	100.0	27	AA512319	DNA encoding class
27	15	100.0	27	AA512320	DNA encoding class
28	15	100.0	27	AA512321	DNA encoding class
29	15	100.0	27	AA512322	DNA encoding class
30	15	100.0	27	AA512323	DNA encoding class
31	15	100.0	27	AA512324	DNA encoding class
32	15	100.0	27	AA512325	DNA encoding class
33	15	100.0	27	AA512326	DNA encoding class
34	15	100.0	27	AA512327	DNA encoding class
35	15	100.0	27	AA512328	DNA encoding class
36	15	100.0	27	AA512329	DNA encoding class
37	15	100.0	27	AA512330	DNA encoding class
38	15	100.0	27	AA512331	DNA encoding class
39	15	100.0	27	AA512332	DNA encoding class
40	15	100.0	27	AA512333	DNA encoding class
41	15	100.0	27	AA512334	DNA encoding class
42	15	100.0	27	AA512335	DNA encoding class
43	15	100.0	27	AA512336	DNA encoding class
44	15	100.0	27	AA512337	DNA encoding class
45	15	100.0	27	AA512404	DNA encoding class

ALIGNMENTS

RESULT 1
AA512347 standard; DNA; 15 BP.
ID AA512347;
AC AA512347;
XX
XX
XX 21-NOV-2001 (first entry)
XX
XX
XX DNA encoding deoxyribozyme #7.
XX
XX Deoxyribozyme; cytosilatic; endonuclease: RNA cleavage; DNA cleavage;
XX gene therapy; plant; fungus; bacteria; mammal; ribozyme; ss.
XX
XX Synthetic.
XX
XX WO200159102-A2.
XX
XX 16-AUG-2001.
XX
XX
XX 08-FEB-2001; 2001WO-US04223.
XX
XX
XX 08-FEB-2000; 2000US-0181360.
XX
XX 31-MAR-2000; 2000US-0193646.
XX
XX (RIBO-) RIBOZYME PHARM INC.
XX (UYTA) UNIV YALE.
XX
XX Breaker R, Belgelman L, Emilsson G;
XX WPI; 2001-536526/59.
XX
XX New nucleic acids with endonuclease activity, such as ribozymes and
XX PT


```

PT nucleozymes, for modulating gene expression in a plant, mammalian,
PI bacterial or fungal cell -
PS
PS Claim 49; Page 77; 96pp; English.
XX
CC The invention relates to nucleic acid molecules with endonuclease
CC activity, which are particularly useful for cleavage of RNA or DNA.
CC The nucleic acids are used in a pharmaceutical composition and are used
CC to modulate expression of a gene in a plant, mammalian, bacterial or
CC fungal cell. They are used to cleave a separate nucleic acid, preferably
CC RNA. The nucleic acids are used to inhibit gene expression and/or cell
CC proliferation, and can be used to treat a disease or condition. More
CC than one nucleic acid can be independently targeted to the same or
CC different sites in a cell. The nucleic acids may be used to study DNA.
CC The modifications to the nucleic acids optimises their catalytic activity
CC and can maintain or enhance their activity. They exhibit a high degree
CC of specificity for RNA. The present sequence represents the coding
CC sequence of deoxyribozyme #7 used in the method of the invention.
CC
XX Sequence 15 BP; 7 A; 1 C; 4 G; 3 U; 0 other;
XX
QY 1 agaauacgugaagau 15
Db 1 agaauacgugaagau 15
XX
RESULT 2
AA512295
ID AA512295 standard; DNA; 27 BP.
AC
XX AA512295;
XX
DT 21-NOV-2001 (first entry)
XX
DE DNA encoding class V ribozyme #7.
XX
KM Ribozyme; cytostatic; endonuclease; RNA cleavage; DNA cleavage;
KM gene therapy; plant; fungus; bacteria; mammal; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..4
FT /*tag= a
FT /mod_base= OTHER
FT /note= "OTHER = 2'-O-methyl nucleotides"
FT modified_base 23..27
FT /*tag= b
FT /mod_base= OTHER
FT /note= "OTHER = 2'-O-methyl nucleotides"
FT modified_base 27
FT /*tag= c
FT /mod_base= c
FT /note= "3',3'-inverted deoxyabasic motety"
XX
PD 16-ADG-2001.
XX
XX 08-FEB-2001; 2001WO-US04223.
XX PF
XX 08-FEB-2000; 2000US-0181360.
XX PR
XX 31-MAR-2000; 2000US-0193646.
XX
XX (RIBO-) RIBOZYME PHARM INC.
XX (UYVA) UNIV YALE.
XX
PI Breaker R, Beigelman L, Emilsson G;

```

```

XX  WPI: 2001-536526/59.
DR
XX  New nucleic acids with endonuclease activity, such as ribozymes and
PT  nucleozymes, for modulating gene expression in a plant, mammalian,
PT  bacterial or fungal cell
PS
XX  Example 1; Page 71; 96pp; English.
PS
XX  The invention relates to nucleic acid molecules with endonuclease
CC  activity, which are particularly useful for cleavage of RNA or DNA.
CC  The nucleic acids are used in a pharmaceutical composition and are used
CC  to modulate expression of a gene in a plant, mammalian, bacterial or
CC  fungal cell. They are used to cleave a separate nucleic acid, preferably
CC  RNA. The nucleic acids are used to inhibit gene expression and/or cell
CC  proliferation, and can be used to treat a disease or condition. More
CC  than one nucleic acid can be used independently targeted to the same or
CC  different sites in a cell. The nucleic acids may be used to study DNA.
CC  The modifications to the nucleic acids optimises their catalytic activity
CC  and can maintain or enhance their activity. They exhibit a high degree
CC  of specificity for RNA. The present sequence represents the coding
CC  sequence of class V ribozyme #7 used in the method of the invention.
XX
SQ  Sequence 27 BP; 11 A; 3 C; 9 G; 4 U; 0 other;

Query Match          100.0%; Score 15; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 5,5;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

OY      1 agaanaacgugaagau 15
        |||||||
DB      7 agaanaacgugaagau 21

RESULT  3
AAS12296
ID      AAS12296 standard; DNA; 27 BP.
XX
XX      AAS12296;
AC
XX      21-NOV-2001 (first entry)
DT
DE      DNA encoding class V ribozyme #8.
XX
XX      Ribozyme; cytosolic; endonuclease; RNA cleavage; DNA cleavage;
KM      gene therapy; plant; fungus; bacteria; mammal; ss.
XX
XX      Synthetic.
OS
XX
FH      Key
FH      modified_base
PT      1.6
PT      Location/Qualifiers
FT      /*tag= a
FT      /mod_base= OTHER
FT      /note= "OTHER = 2'-O-methyl nucleotides"
FT      20..27
FT      /*tag= b
FT      /mod_base= OTHER
FT      /note= "OTHER = 2'-O-methyl nucleotides"
FT      27
FT      /*tag= c
FT      /mod_base= c
FT      /note= "3',3'-inverted deoxyabasic moiety"
XX
XX      WO200159102-A2.
XX
XX      16-AUG-2001.
XX
XX      08-FEB-2001; 2001WO-US04223.
XX
XX      08-FEB-2000; 2000US-0181360.
XX      31-MAR-2000; 2000US-0193646.
XX

```

PA (RIBO-) RIBOZYME PHARM INC.
PA (UYVA) UNIV VALE.
XX
PI Breaker R, Beigelman L, Emilsson G;
XX
DR WPI; 2001-536526/59.
XX
PT New nucleic acids with endonuclease activity, such as ribozymes and
PT nucleozymes, for modulating gene expression in a plant, mammalian,
PT bacterial or fungal cell
XX
XX
PS Example 1; Page 71; 96pp; English.
XX
CC The invention relates to nucleic acid molecules with endonuclease
CC activity, which are particularly useful for cleavage of RNA or DNA.
CC The nucleic acids are used in a pharmaceutical composition and are used
CC to modulate expression of a gene in a plant, mammalian, bacterial or
CC fungal cell. They are used to cleave a separate nucleic acid, preferably
CC RNA. The nucleic acids are used to inhibit gene expression and/or cell
CC proliferation, and can be used to treat a disease or condition. More
CC than one nucleic acid can be independently targeted to the same or
CC different sites in a cell. The nucleic acids may be used to study DNA.
CC The modifications to the nucleic acids optimises their catalytic activity
CC and can maintain or enhance their activity. They exhibit a high degree
CC of specificity for RNA. The present sequence represents the coding
CC sequence of class V ribozyme #8 used in the method of the invention.
XX
SQ Sequence 27 BP; 11 A; 3 C; 9 G; 4 U; 0 other;

Query Match 100.0%; Score 15; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaagau 15
Db 7 agauaacgugaagau 21

RESULT 4
AAS12297
ID AAS12297 standard; DNA; 27 BP.
XX
AC AAS12297;
XX
DT 21-NOV-2001 (first entry)
XX
DE DNA encoding class V ribozyme #9.
XX
KW Ribozyme; cytosstatic; endonuclease; RNA cleavage; DNA cleavage;
KW gene therapy; plant; fungus; bacteria; mammal; ss.
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT modified_base 1..6
FT /tag- a
FT /mod_base= OTHER
FT /note= "OTHER = 2'-O-methyl nucleotides"
FT 21.27
FT /tag- b
FT /mod_base= OTHER
FT /note= "OTHER = 2'-O-methyl nucleotides"
FT 27
FT /tag- c
FT /mod_base= c
FT /note= "3',3'-inverted deoxyabasic moiety"
XX
XX
PN WO200159102-A2.
XX
PD 16-NOV-2001.
XX
PF 08-FEB-2001; 2001WO-US04223.

XX
PR 08-FEB-2000; 2000US-0181360.
PR 31-MAR-2000; 2000US-0193646.
XX
XX
PA (RIBO-) RIBOZYME PHARM INC.
PA (UYVA) UNIV VALE.
XX
XX
PI Breaker R, Beigelman L, Emilsson G;
XX
DR WPI; 2001-536526/59.
XX
PT New nucleic acids with endonuclease activity, such as ribozymes and
PT nucleozymes, for modulating gene expression in a plant, mammalian,
PT bacterial or fungal cell
XX
XX
PS Example 1; Page 71; 96pp; English.
XX
CC The invention relates to nucleic acid molecules with endonuclease
CC activity, which are particularly useful for cleavage of RNA or DNA.
CC The nucleic acids are used in a pharmaceutical composition and are used
CC to modulate expression of a gene in a plant, mammalian, bacterial or
CC fungal cell. They are used to cleave a separate nucleic acid, preferably
CC RNA. The nucleic acids are used to inhibit gene expression and/or cell
CC proliferation, and can be used to treat a disease or condition. More
CC than one nucleic acid can be independently targeted to the same or
CC different sites in a cell. The nucleic acids may be used to study DNA.
CC The modifications to the nucleic acids optimises their catalytic activity
CC and can maintain or enhance their activity. They exhibit a high degree
CC of specificity for RNA. The present sequence represents the coding
CC sequence of class V ribozyme #9 used in the method of the invention.
XX
SQ Sequence 27 BP; 11 A; 3 C; 9 G; 4 U; 0 other;

Query Match 100.0%; Score 15; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaagau 15
Db 7 agauaacgugaagau 21

RESULT 5
AAS12298
ID AAS12298 standard; DNA; 27 BP.
XX
AC AAS12298;
XX
DT 21-NOV-2001 (first entry)
XX
DE DNA encoding class V ribozyme #10.
XX
KW Ribozyme; cytosstatic; endonuclease; RNA cleavage; DNA cleavage;
KW gene therapy; plant; fungus; bacteria; mammal; ss.
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT modified_base 1..6
FT /tag- a
FT /mod_base= OTHER
FT /note= "OTHER = 2'-O-methyl nucleotides"
FT 12
FT /tag- b
FT /mod_base= OTHER
FT /note= "OTHER = 2'-O-methyl nucleotide"
FT 21.27
FT /tag- c
FT /mod_base= OTHER
FT /note= "OTHER = 2'-O-methyl nucleotides"
FT 27
FT /tag- d
FT modified_base

```

FT FT /mod_base= C
FT FT /note= "3',3'-inverted deoxyabasic moiety"
XX XX
PN MO200159102-A2.
XX PD 16-AUG-2001.
XX PF 08-FEB-2001; 2001WO-US04223.
XX PR 08-FEB-2000; 2000US-0181360.
XX PR 31-MAR-2000; 2000US-0193646.
XX XX (RIBO-) RIBOZYME PHARM INC.
PA (UYVA ) UNIV YALE.
XX XX
PI Breaker R, Belgelman L, Emilsson G;
XX WPI; 2001-536526/59.
XX DR
XX XX New nucleic acids with endonuclease activity, such as ribozymes and
PT nucleozymes, for modulating gene expression in a plant, mammalian,
PT bacterial or fungal cell -
XX XX
XX PS Example 1; Page 71; 96pp; English.
XX CC The invention relates to nucleic acid molecules with endonuclease
CC activity, which are particularly useful for cleavage of RNA or DNA.
CC CC The nucleic acids are used in a pharmaceutical composition and are used
CC to modulate expression of a gene in a plant, mammalian, bacterial or
CC fungal cell. They are used to cleave a separate nucleic acid, preferably
CC RNA. The nucleic acids are used to inhibit gene expression and/or cell
CC proliferation, and can be used to treat a disease or condition. More
CC than one nucleic acid can be independently targeted to the same or
CC different sites in a cell. The nucleic acids may be used to study DNA.
CC The modifications to the nucleic acids optimises their catalytic activity
CC and can maintain or enhance their activity. They exhibit a high degree
CC of specificity for RNA. The present sequence represents the coding
CC sequence of class V ribozyme #10 used in the method of the invention.
XX XX
SQ Sequence 27 BP; 11 A; 3 C; 9 G; 4 U; 0 other;

Query Match 100.0%; Score 15; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaagau 15
Db 7 agauaacgugaagau 21

RESULT 6
AAS12299
ID AAS12299 standard; DNA; 27 BP.
XX
AC AAS12299;
XX
DT 21-NOV-2001 (first entry)
XX
DE DNA encoding class V ribozyme #11.
XX
KM Ribozyme; cytosstatic; endonuclease; RNA cleavage; DNA cleavage;
KM gene therapy; plant; fungus; bacteria; mammal; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1.7
FT /tag= a
FT /mod_base= OTHER
FT /note= "OTHER = 2'-O-methyl nucleotides"
FT modified_base 21..27
FT /tag= b

```

```

FT FT /mod_base= OTHER
FT FT /note= "OTHER = 2'-O-methyl nucleotides"
XX XX
PN modified_base 27
XX FT /tag= c
FT FT /mod_base= C
FT FT /note= "3',3'-inverted deoxyabasic moiety"
XX XX
PN MO200159102-A2.
XX PD 16-AUG-2001.
XX PF 08-FEB-2001; 2001WO-US04223.
XX PR 08-FEB-2000; 2000US-0181360.
XX PR 31-MAR-2000; 2000US-0193646.
XX XX (RIBO-) RIBOZYME PHARM INC.
PA (UYVA ) UNIV YALE.
XX XX
PI Breaker R, Belgelman L, Emilsson G;
XX WPI; 2001-536526/59.
XX DR
XX XX New nucleic acids with endonuclease activity, such as ribozymes and
PT nucleozymes, for modulating gene expression in a plant, mammalian,
PT bacterial or fungal cell -
XX XX
XX PS Example 1; Page 71; 96pp; English.
XX CC The invention relates to nucleic acid molecules with endonuclease
CC activity, which are particularly useful for cleavage of RNA or DNA.
CC CC The nucleic acids are used in a pharmaceutical composition and are used
CC to modulate expression of a gene in a plant, mammalian, bacterial or
CC fungal cell. They are used to cleave a separate nucleic acid, preferably
CC RNA. The nucleic acids are used to inhibit gene expression and/or cell
CC proliferation, and can be used to treat a disease or condition. More
CC than one nucleic acid can be independently targeted to the same or
CC different sites in a cell. The nucleic acids may be used to study DNA.
CC The modifications to the nucleic acids optimises their catalytic activity
CC and can maintain or enhance their activity. They exhibit a high degree
CC of specificity for RNA. The present sequence represents the coding
CC sequence of class V ribozyme #11 used in the method of the invention.
XX XX
SQ Sequence 27 BP; 11 A; 3 C; 9 G; 4 U; 0 other;

Query Match 100.0%; Score 15; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaagau 15
Db 7 agauaacgugaagau 21

RESULT 7
AAS12300
ID AAS12300 standard; DNA; 27 BP.
XX
AC AAS12300;
XX
DT 21-NOV-2001 (first entry)
XX
DE DNA encoding class V ribozyme #12.
XX
KM Ribozyme; cytosstatic; endonuclease; RNA cleavage; DNA cleavage;
KM gene therapy; plant; fungus; bacteria; mammal; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1.6
FT /tag= a

```

FT	/mod_base= OTHER
FT	/note= "OTHER = 2'-O-methyl nucleotides"
FT	21..27
FT	/*tag= b
FT	/mod_base= OTHER
FT	/note= "OTHER = 2'-O-methyl nucleotides"
FT	27
FT	/*tag= C
FT	/mod_base= C
FT	/note= "3',3'-Inverted deoxybasic moiety"
PN	WO200159102-A2.
XX	16-AUG-2001.
XX	08-FEB-2001; 2001WO-US04223.
XX	08-FEB-2000; 2000US-0181360.
PR	31-MAR-2000; 2000US-0193646.
XX	(RIBO-) RIBOZYME PHARM INC.
PA	(UYA) UNIV YALE.
XX	Breaker R, Beigelman L, Emilsson G;
PI	WPI; 2001-536526/59.
XX	New nucleic acids with endonuclease activity, such as ribozymes and
PT	nucleozymes, for modulating gene expression in a plant, mammalian,
PT	bacterial or fungal cell -
XX	Example 1; Page 71; 96pp; English.
PS	The invention relates to nucleic acid molecules with endonuclease
CC	activity, which are particularly useful for cleavage of RNA or DNA.
CC	The nucleic acids are used in a pharmaceutical composition and are used
CC	to modulate expression of a gene in a plant, mammalian, bacterial or
CC	fungal cell. They are used to cleave a separate nucleic acid, preferably
CC	RNA. The nucleic acids are used to inhibit gene expression and/or cell
CC	proliferation, and can be used to treat a disease or condition. More
CC	than one nucleic acid can be independently targeted to the same or
CC	different sites in a cell. The nucleic acids may be used to study DNA.
CC	The modifications to the nucleic acids optimises their catalytic activity
CC	and can maintain or enhance their activity. They exhibit a high degree
CC	of specificity for RNA. The present sequence represents the coding
CC	sequence of class V ribozyme #12 used in the method of the invention.
XX	
SO	Sequence 27 BP; 11 A; 3 C; 9 G; 4 U; 0 other;
Query Match	100.0%; Score 15; DB 22; Length 27;
Best Local Similarity	100.0%; Pred. No. 5.5;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 agaunaacugagaau 15
Db	7 agaunaacugagaau 21
RESULT	8
AA512301	
ID	AA512301 standard; DNA; 27 BP.
XX	
AC	AA512301;
XX	
DT	21-NOV-2001 (first entry)
XX	
DE	DNA encoding class V ribozyme #13.
XX	
KW	Ribozyme; cytosolic; endonuclease; RNA cleavage; DNA cleavage;
KW	gene therapy; plant; fungus; bacteria; mammal; ss.
XX	
SS	Synthetic.

[illegible]

XX	21-NOV-2001 (first entry)
DE	DNA encoding class V ribozyme #14.
KM	Ribozyme; cytosolic; endonuclease; RNA cleavage; DNA cleavage;
KM	gene therapy; plant; fungus; bacteria; mammal; ss.
XX	
OS	Synthetic.
XX	
XX	Key
FT	Location/Qualifiers
FT	1..6
FT	/*tag= a
FT	/mod_base= OTHER
FT	/note= "OTHER = 2'-O-methyl nucleotides"
FT	17
FT	/*tag= b
FT	/mod_base= a
FT	/note= "OTHER = 2'-O-methyl nucleotide"
FT	21..27
FT	/*tag= c
FT	/mod_base= OTHER
FT	/note= "OTHER = 2'-O-methyl nucleotides"
FT	27
FT	/*tag= d
FT	/mod_base= c
FT	/note= "3',3'-inverted deoxybasic moiety"
XX	
PN	WO200159102-A2.
PD	16-AUG-2001.
XX	
PE	08-FEB-2001; 2001WO-US04223.
XX	
PR	08-FEB-2000; 2000US-0181360.
PR	31-MAR-2000; 2000US-0193646.
XX	
PA	(RIBO-) RIBOZYME PHARM INC.
PA	(UTYA) UNIT YALE.
XX	
PI	Breaker R, Belgelman L, Emlisson G;
XX	
DR	WPI, 2001-536526/59.
XX	
PT	New nucleic acids with endonuclease activity, such as ribozymes and
PT	nucleozymes, for modulating gene expression in a plant, mammalian,
PT	bacterial or fungal cell
XX	
PS	Example 1; Page 71; 96pp; English.
XX	
CC	The invention relates to nucleic acid molecules with endonuclease
CC	activity, which are particularly useful for cleavage of RNA or DNA.
CC	The nucleic acids are used in a pharmaceutical composition and are used
CC	to modulate expression of a gene in a plant, mammalian, bacterial or
CC	fungal cell. They are used to cleave a separate nucleic acid, preferably
CC	RNA. The nucleic acids are used to inhibit gene expression and/or cell
CC	proliferation, and can be used to treat a disease or condition. More
CC	than one nucleic acid can be independently targeted to the same or
CC	different sites in a cell. The nucleic acids may be used to study DNA.
CC	The modifications to the nucleic acids optimises their catalytic activity
CC	and can maintain or enhance their activity. They exhibit a high degree
CC	of specificity for RNA. The present sequence represents the coding
CC	sequence of class V ribozyme #14 used in the method of the invention.
XX	
SO	Sequence 27 BP; 11 A; 3 C; 9 G; 4 U; 0 other;
XX	
QY	Query Match 100.0%; Score 15; DB 22; Length 27;
XX	Best Local Similarity 100.0%; Pred. No. 5.5;
DB	Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB	1 aguaaagcugagaagau 15
DB	
DB	7 aguaaagcugagaagau 21

CC	The invention relates to nucleic acid molecules with endonuclease activity, which are particularly useful for cleavage of RNA or DNA.
CC	The nucleic acids are used in a pharmaceutical composition and are used to modulate expression of a gene in a plant, mammalian, bacterial or fungal cell.
CC	RNA. The nucleic acids are used to inhibit gene expression and/or cell proliferation, and can be used to treat a disease or condition. More than one nucleic acid can be independently targeted to the same or different sites in a cell. The nucleic acids may be used to study DNA.
CC	The modifications to the nucleic acids optimises their catalytic activity and can maintain or enhance their activity. They exhibit a high degree of specificity for RNA. The present sequence represents the coding
XX	Key
XX	Synthetic.
XX	DNA encoding class V ribozyme #15.
XX	Ribozyme; cytosolic; endonuclease; RNA cleavage; DNA cleavage;
XX	gene therapy; plant; fungus; bacteria; mammal; ss.
XX	Location/Qualifiers
XX	1..6
XX	/tag= a
XX	/mod_base= OTHER
XX	/note= "OTHER = 2'-O-methyl nucleotides"
XX	11
XX	/tag= b
XX	/mod_base= a
XX	/note= "OTHER =-2'-O-methyl nucleotide"
XX	16
XX	/tag= c
XX	/mod_base= g
XX	/note= "OTHER = 2'-O-methyl nucleotide"
XX	21..27
XX	/tag= d
XX	/mod_base= OTHER
XX	/note= "OTHER = 2'-O-methyl nucleotides"
XX	27
XX	/tag= e
XX	/mod_base= c
XX	/note= "3',3'-inverted deoxyabasic moiety"
XX	WO200159102-A2.
XX	16-AUG-2001.
XX	08-FEB-2001; 2001WO-US04223.
XX	08-FEB-2000; 2000US-0181360.
XX	31-MAR-2000; 2000US-0193646.
XX	(RIBO-) RIBOZYME PHARM INC.
XX	(UYTA) UNTV YALE.
XX	Breaker R, Beigelman L, Emilsson G;
XX	MPI; 2001-536526/59.
XX	New nucleic acids with endonuclease activity, such as ribozymes and nucleozymes, for modulating gene expression in a plant, mammalian, bacterial or fungal cell
XX	Example 1; Page 71; 96pp: English.

CC sequence of class V ribozyme #15 used in the method of the invention.
XX Sequence 27 BP; 11 A; 3 C; 9 G; 4 U; 0 other;
SQ

Query Match 100.0%; Score 15; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agaanaacgugaagau 15
Db 7 agaanaacgugaagau 21

RESULT 11
AAS12304
ID AAS12304 standard; DNA; 27 BP.
XX AAS12304;
AC
XX 21-NOV-2001 (first entry)
DT
XX DNA encoding class V ribozyme #16.
DE
XX Ribozyme; cytosolic; endonuclease; RNA cleavage; DNA cleavage;
KM gene therapy; plant; fungus; bacteria; mammal; ss.
XX Synthetic.
OS
XX
FH Key Location/Qualifiers
FT 1..6
FT modified_base
FT /*tag= a
FT /mod_base= OTHER
FT /note= "OTHER = 2'-O-methyl nucleotides"
FT 21..27
FT /*tag= b
FT /mod_base= OTHER
FT /note= "OTHER = 2'-O-methyl nucleotides"
FT 27
FT /*tag= c
FT /mod_base= c
FT /note= "3',3'-inverted deoxyabasic moiety"
PN WO200159102-A2.
XX 16-AUG-2001.
PD
XX 08-FEB-2001; 2001WO-US04223.
PF
XX 08-FEB-2000; 2000US-0181360.
PR 31-MAR-2000; 2000US-0193646.
XX
XX (RIBO-) RIBOZYME PHARM INC.
PA (UYVA) UNIV YALE.
XX
XX Breaker R, Beigelman L, Emilsson G;
PI WPI; 2001-536526/59.
DR
XX New nucleic acids with endonuclease activity, such as ribozymes and
PT nucleozymes, for modulating gene expression in a plant, mammalian,
PT bacterial or fungal cell
XX
XX Example 1; Page 71; 96pp; English.
PS
XX The invention relates to nucleic acid molecules with endonuclease
CC activity, which are particularly useful for cleavage of RNA or DNA.
CC The nucleic acids are used in a pharmaceutical composition and are used
CC to modulate expression of a gene in a plant, mammalian, bacterial or
CC fungal cell. They are used to cleave a separate nucleic acid, preferably
CC RNA. The nucleic acids are used to inhibit gene expression and/or cell
CC proliferation, and can be used to treat a disease or condition. More
CC than one nucleic acid can be independently targeted to the same or

CC different sites in a cell. The nucleic acids may be used to study DNA.
CC The modifications to the nucleic acids optimises their catalytic activity
CC and can maintain or enhance their activity. They exhibit a high degree
CC of specificity for RNA. The present sequence represents the coding
CC sequence of class V ribozyme #16 used in the method of the invention.
XX
SQ Sequence 27 BP; 11 A; 3 C; 9 G; 4 U; 0 other;

Query Match 100.0%; Score 15; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agaanaacgugaagau 15
Db 7 agaanaacgugaagau 21

RESULT 12
AAS12305
ID AAS12305 standard; DNA; 27 BP.
XX AAS12305;
AC
XX 21-NOV-2001 (first entry)
DT
XX DNA encoding class V ribozyme #17.
DE
XX Ribozyme; cytosolic; endonuclease; RNA cleavage; DNA cleavage;
KM gene therapy; plant; fungus; bacteria; mammal; ss.
XX Synthetic.
OS
XX
FH Key Location/Qualifiers
FT 1..6
FT modified_base
FT /*tag= a
FT /mod_base= OTHER
FT /note= "OTHER = 2'-O-methyl nucleotides"
FT 19
FT modified_base
FT /*tag= b
FT /mod_base= g
FT /note= "OTHER = 2'-O-methyl nucleotide"
FT 21..27
FT /*tag= c
FT /mod_base= OTHER
FT /note= "OTHER = 2'-O-methyl nucleotides"
FT 27
FT /*tag= d
FT /mod_base= c
FT /note= "3',3'-inverted deoxyabasic moiety"
PN WO200159102-A2.
XX 16-AUG-2001.
PD
XX 08-FEB-2001; 2001WO-US04223.
PF
XX 08-FEB-2000; 2000US-0181360.
PR 31-MAR-2000; 2000US-0193646.
XX
XX (RIBO-) RIBOZYME PHARM INC.
PA (UYVA) UNIV YALE.
XX
XX Breaker R, Beigelman L, Emilsson G;
PI WPI; 2001-536526/59.
DR
XX New nucleic acids with endonuclease activity, such as ribozymes and
PT nucleozymes, for modulating gene expression in a plant, mammalian,
PT bacterial or fungal cell
XX
XX Example 1; Page 71; 96pp; English.
PS


```
PR 08-FEB-2000; 2000US-0181360.
PR 31-MAR-2000; 2000US-0193646.
XX
XX (RIBO-) RIBOZYME PHARM INC.
XX (UYA ) UNIV YALE.
XX
XX Breaker R, Belgelman L, Emilsson G;
XX WPI; 2001-536526/59.
XX
XX New nucleic acids with endonuclease activity, such as ribozymes and
XX PT nucleozymes, for modulating gene expression in a plant, mammalian,
XX PT bacterial or fungal cell
XX
XX Example 1; Page 71; 96pp; English.
XX
XX The invention relates to nucleic acid molecules with endonuclease
XX CC activity, which are particularly useful for cleavage of RNA or DNA.
XX CC The nucleic acids are used in a pharmaceutical composition and are used
XX CC to modulate expression of a gene in a plant, mammalian, bacterial or
XX CC fungal cell. They are used to cleave a separate nucleic acid, preferably
XX CC RNA. The nucleic acids are used to inhibit gene expression and/or cell
XX CC proliferation, and can be used to treat a disease or condition. More
XX CC than one nucleic acid can be independently targeted to the same or
XX CC different sites in a cell. The nucleic acids may be used to study DNA.
XX CC The modifications to the nucleic acids optimises their catalytic activity
XX CC and can maintain or enhance their activity. They exhibit a high degree
XX CC of specificity for RNA. The present sequence represents the coding
XX CC sequence of class V ribozyme #19 used in the method of the invention.
XX
XX Sequence 27 BP; 11 A; 3 C; 9 G; 4 U; 0 other;
SO

Query Match 100.0%; Score 15; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaagan 15
    |||||
Db 7 agauaacgugaagan 21

RESULT 15
AAS12308
ID AAS12308 standard; DNA; 27 BP.
XX
AC AAS12308;
XX
DT 21-NOV-2001 (first entry)
XX
DE DNA encoding class V ribozyme #20.
XX
XX Ribozyme; cytosolic; endonuclease; RNA cleavage; DNA cleavage;
XX KM gene therapy; plant; fungus; bacteria; mammal; ss.
XX OS Synthetic.
XX
XX Key Location/Qualifiers
XX FH modified_base 1..6
XX FT /*tag- a
XX FT /mod_base- OTHER
XX FT /note- "OTHER = 2'-O-methyl nucleotides"
XX FT 21..27
XX FT /*tag- b
XX FT /mod_base- OTHER
XX FT /note- "OTHER = 2'-O-methyl nucleotides"
XX FT 27
XX FT /*tag- c
XX FT /mod_base- c
XX FT /note- "3',3'-inverted deoxybasic moiety"
XX PN
XX WO200159102-A2.
```

```
PD 16-AUG-2001.
XX
XX PF 08-FEB-2001; 2001WO-US04223.
XX PF
XX PR 08-FEB-2000; 2000US-0181360.
XX PR 31-MAR-2000; 2000US-0193646.
XX
XX (RIBO-) RIBOZYME PHARM INC.
XX PA (UYA ) UNIV YALE.
XX
XX Breaker R, Belgelman L, Emilsson G;
XX WPI; 2001-536526/59.
XX
XX New nucleic acids with endonuclease activity, such as ribozymes and
XX PT nucleozymes, for modulating gene expression in a plant, mammalian,
XX PT bacterial or fungal cell
XX
XX Example 1; Page 71; 96pp; English.
XX
XX The invention relates to nucleic acid molecules with endonuclease
XX CC activity, which are particularly useful for cleavage of RNA or DNA.
XX CC The nucleic acids are used in a pharmaceutical composition and are used
XX CC to modulate expression of a gene in a plant, mammalian, bacterial or
XX CC fungal cell. They are used to cleave a separate nucleic acid, preferably
XX CC RNA. The nucleic acids are used to inhibit gene expression and/or cell
XX CC proliferation, and can be used to treat a disease or condition. More
XX CC than one nucleic acid can be independently targeted to the same or
XX CC different sites in a cell. The nucleic acids may be used to study DNA.
XX CC The modifications to the nucleic acids optimises their catalytic activity
XX CC and can maintain or enhance their activity. They exhibit a high degree
XX CC of specificity for RNA. The present sequence represents the coding
XX CC sequence of class V ribozyme #20 used in the method of the invention.
XX
XX Sequence 27 BP; 11 A; 3 C; 9 G; 4 U; 0 other;
SO

Query Match 100.0%; Score 15; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaagan 15
    |||||
Db 7 agauaacgugaagan 21
```

Search completed: September 9, 2002, 01:51:27
Job time: 4527 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 8, 2002, 23:03:20 : Search time 1931.22 seconds

(without alignments)
162.539 Million cell updates/sec

Title: US-09-780-929-97

Sequence: 15
1 agaaacgugaaga 15

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Genembl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_of:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Match	Length	DB	ID	Description
--------	-------	-------	--------	----	----	-------------

1	15	100.0	15	6	AX214295	AX214295 Sequence
2	15	100.0	27	6	AX214239	AX214239 Sequence
3	15	100.0	28	6	AX214237	AX214237 Sequence
4	15	100.0	28	6	AX214238	AX214238 Sequence
5	15	100.0	28	6	AX214240	AX214240 Sequence
6	15	100.0	28	6	AX214241	AX214241 Sequence
7	15	100.0	28	6	AX214242	AX214242 Sequence
8	15	100.0	28	6	AX214243	AX214243 Sequence
9	15	100.0	28	6	AX214244	AX214244 Sequence
10	15	100.0	28	6	AX214245	AX214245 Sequence
11	15	100.0	28	6	AX214246	AX214246 Sequence
12	15	100.0	28	6	AX214247	AX214247 Sequence
13	15	100.0	28	6	AX214248	AX214248 Sequence
14	15	100.0	28	6	AX214249	AX214249 Sequence
15	15	100.0	28	6	AX214250	AX214250 Sequence
16	15	100.0	28	6	AX214251	AX214251 Sequence
17	15	100.0	28	6	AX214252	AX214252 Sequence
18	15	100.0	28	6	AX214253	AX214253 Sequence
19	15	100.0	28	6	AX214254	AX214254 Sequence
20	15	100.0	28	6	AX214255	AX214255 Sequence
21	15	100.0	28	6	AX214256	AX214256 Sequence
22	15	100.0	28	6	AX214257	AX214257 Sequence
23	15	100.0	28	6	AX214258	AX214258 Sequence
24	15	100.0	28	6	AX214259	AX214259 Sequence
25	15	100.0	28	6	AX214260	AX214260 Sequence
26	15	100.0	28	6	AX214261	AX214261 Sequence
27	15	100.0	28	6	AX214262	AX214262 Sequence
28	15	100.0	28	6	AX214263	AX214263 Sequence
29	15	100.0	28	6	AX214264	AX214264 Sequence
30	15	100.0	28	6	AX214265	AX214265 Sequence
31	15	100.0	28	6	AX214266	AX214266 Sequence
32	15	100.0	28	6	AX214267	AX214267 Sequence
33	15	100.0	28	6	AX214268	AX214268 Sequence
34	15	100.0	28	6	AX214269	AX214269 Sequence
35	15	100.0	28	6	AX214270	AX214270 Sequence
36	15	100.0	28	6	AX214271	AX214271 Sequence
37	15	100.0	28	6	AX214272	AX214272 Sequence
38	15	100.0	28	6	AX214273	AX214273 Sequence
39	15	100.0	28	6	AX214274	AX214274 Sequence
40	15	100.0	28	6	AX214275	AX214275 Sequence
41	15	100.0	28	6	AX214276	AX214276 Sequence
42	15	100.0	28	6	AX214277	AX214277 Sequence
43	15	100.0	28	6	AX214278	AX214278 Sequence
44	15	100.0	28	6	AX214279	AX214279 Sequence
45	15	100.0	28	6	AX214280	AX214280 Sequence

ALIGNMENTS

RESULT 1
AX214295 LOCUS AX214295 15 bp mRNA
DEFINITION Sequence 108 from Patent WO0159102. linear PAT 06-SEP-2001
ACCESSION AX214295
VERSION AX214295.1 GI:15524372
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
synthetic construct.
artificial sequence.
REFERENCE
1 (bases 1 to 15)
AUTHORS Breaker, R. and Emlison, G.
TITLE Nucleozymes with endonuclease activity
JOURNAL Patent: WO 0159102-A 108 16-AUG-2001;
RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)
FEATURES
source location/Qualifiers
1..15
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Nucleic Acid"

BASE COUNT
ORIGIN

7 a 1 c 4 g 3 t

Query Match 100.0%; Score 15; DB 6; Length 15;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 agaaacgugaagau 15
|||||:|||||
DB 1 AGATAACGTGAGAT 15

RESULT 2

AX214239

LOCUS AX214239 27 bp mRNA linear PAT 06-SEP-2001

DEFINITION Sequence 52 from Patent WO0159102.

ACCESSION AX214239

VERSION AX214239.1 GI:15524316

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 27)

AUTHORS Breaker,R. and Emlisson,G.

TITLE Nucleozymes with endonuclease activity

JOURNAL Patent: WO 0159102-A 52 16-AUG-2001;

RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)

FEATURES Location/Qualifiers

source

1..27 /organism="synthetic construct"

/db_xref="taxon:32630"

/note="Nucleic Acid"

BASE COUNT 11 a 3 c 9 g 4 t

ORIGIN

Query Match 100.0%; Score 15; DB 6; Length 27;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 agaaacgugaagau 15
|||||:|||||
DB 7 AGATAACGTGAGAT 21

RESULT 3

AX214237

LOCUS AX214237 28 bp mRNA linear PAT 06-SEP-2001

DEFINITION Sequence 50 from Patent WO0159102.

ACCESSION AX214237

VERSION AX214237.1 GI:15524314

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 28)

AUTHORS Breaker,R. and Emlisson,G.

TITLE Nucleozymes with endonuclease activity

JOURNAL Patent: WO 0159102-A 50 16-AUG-2001;

RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)

FEATURES Location/Qualifiers

source

1..28 /organism="synthetic construct"

/db_xref="taxon:32630"

/note="Nucleic Acid"

misc_feature 1..6 /note="2'-O-Methyl"

misc_feature 21..27 /note="2'-O-Methyl"

misc_feature 28 /note="n stands for inverted deoxybasic derivative"

BASE COUNT 11 a 3 c 9 g 4 t 1 others

ORIGIN

Query Match 100.0%; Score 15; DB 6; Length 28;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 agaaacgugaagau 15
|||||:|||||
DB 7 AGATAACGTGAGAT 21

RESULT 4

AX214238

LOCUS AX214238 28 bp mRNA linear PAT 06-SEP-2001

DEFINITION Sequence 51 from Patent WO0159102.

ACCESSION AX214238

VERSION AX214238.1 GI:15524315

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 28)

AUTHORS Breaker,R. and Emlisson,G.

TITLE Nucleozymes with endonuclease activity

JOURNAL Patent: WO 0159102-A 51 16-AUG-2001;

RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)

FEATURES Location/Qualifiers

source

1..28 /organism="synthetic construct"

/db_xref="taxon:32630"

/note="Nucleic Acid"

misc_feature 1..5 /note="2'-O-Methyl"

misc_feature 21..27 /note="2'-O-Methyl"

misc_feature 28 /note="n stands for inverted deoxybasic derivative"

BASE COUNT 11 a 3 c 9 g 4 t 1 others

ORIGIN

Query Match 100.0%; Score 15; DB 6; Length 28;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 agaaacgugaagau 15
|||||:|||||
DB 7 AGATAACGTGAGAT 21

RESULT 5

AX214240

LOCUS AX214240 28 bp mRNA linear PAT 06-SEP-2001

DEFINITION Sequence 53 from Patent WO0159102.

ACCESSION AX214240

VERSION AX214240.1 GI:15524317

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 28)

AUTHORS Breaker,R. and Emlisson,G.

TITLE Nucleozymes with endonuclease activity

JOURNAL Patent: WO 0159102-A 53 16-AUG-2001;

RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)

FEATURES Location/Qualifiers

source

1..28 /organism="synthetic construct"

/db_xref="taxon:32630"

/note="Nucleic Acid"

misc_feature 1..5 /note="2'-O-Methyl"

misc_feature 22..27 /note="2'-O-Methyl"

misc_feature 28 /note="2'-O-Methyl"

BASE COUNT

BASE COUNT 11 a 3 c 9 g 4 t 1 others
ORIGIN

Query Match 100.0%; Score 15; DB 6; Length 28;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 agauaacgugaagan 15
Db 7 AGATAACCTGAAGAT 21

RESULT 6
AX214241 28 bp mRNA linear PAT 06-SEP-2001
LOCUS
DEFINITION Sequence 54 from Patent WO0159102.
ACCESSION AX214241
VERSION AX214241.1 GI:15524318
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 28)
AUTHORS Breaker,R. and Emilsson,G.
TITLE Nucleozymes with endonuclease activity
JOURNAL Patent: WO 0159102-A 54 16-AUG-2001;
RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)

FEATURES
source location/Qualifiers
1..28
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Nucleic Acid"

misc_feature 1..5
/note="2'-O-Methyl"
misc_feature 23..27
/note="2'-O-Methyl"
misc_feature 28
/note="n stands for inverted deoxybasic derivative"
BASE COUNT 11 a 3 c 9 g 4 t 1 others
ORIGIN

Query Match 100.0%; Score 15; DB 6; Length 28;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 agauaacgugaagan 15
Db 7 AGATAACCTGAAGAT 21

RESULT 7
AX214242 28 bp mRNA linear PAT 06-SEP-2001
LOCUS
DEFINITION Sequence 55 from Patent WO0159102.
ACCESSION AX214242
VERSION AX214242.1 GI:15524319
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 28)
AUTHORS Breaker,R. and Emilsson,G.
TITLE Nucleozymes with endonuclease activity
JOURNAL Patent: WO 0159102-A 55 16-AUG-2001;
RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)

FEATURES
source location/Qualifiers
1..28
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Nucleic Acid"

misc_feature 1..4
/note="2'-O-Methyl"
misc_feature 24..27
/note="2'-O-Methyl"
misc_feature 28
/note="n stands for inverted deoxybasic derivative"
BASE COUNT 11 a 3 c 9 g 4 t 1 others
ORIGIN

Query Match 100.0%; Score 15; DB 6; Length 28;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 agauaacgugaagan 15
Db 7 AGATAACCTGAAGAT 21

RESULT 8
AX214243 28 bp mRNA linear PAT 06-SEP-2001
LOCUS
DEFINITION Sequence 56 from Patent WO0159102.
ACCESSION AX214243
VERSION AX214243.1 GI:15524320
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 28)
AUTHORS Breaker,R. and Emilsson,G.
TITLE Nucleozymes with endonuclease activity
JOURNAL Patent: WO 0159102-A 56 16-AUG-2001;
RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)

FEATURES
source location/Qualifiers
1..28
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Nucleic Acid"

misc_feature 1..4
/note="2'-O-Methyl"
misc_feature 23..27
/note="2'-O-Methyl"
misc_feature 28
/note="n stands for inverted deoxybasic derivative"
BASE COUNT 11 a 3 c 9 g 4 t 1 others
ORIGIN

Query Match 100.0%; Score 15; DB 6; Length 28;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 agauaacgugaagan 15
Db 7 AGATAACCTGAAGAT 21

RESULT 9
AX214244 28 bp mRNA linear PAT 06-SEP-2001
LOCUS
DEFINITION Sequence 57 from Patent WO0159102.
ACCESSION AX214244
VERSION AX214244.1 GI:15524321
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 28)
AUTHORS Breaker,R. and Emilsson,G.
TITLE Nucleozymes with endonuclease activity
JOURNAL Patent: WO 0159102-A 57 16-AUG-2001;
RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)

FEATURES
source location/Qualifiers
1..28
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Nucleic Acid"

FEATURES
source
Location/Qualifiers
1.28
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Nucleic Acid"
misc_feature
1.6
/note="2'-O-Methyl"
misc_feature
20..27
/note="2'-O-Methyl"
misc_feature
28
/note="n stands for inverted deoxybasic derivative"
BASE COUNT
11 a 3 c 9 g 4 t 1 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 15; DB 6; Length 28;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaagu 15
|||||:|||||:
Db 7 AGATACGTGAGAT 21

RESULT 10
AX214245 28 bp mRNA linear PAT 06-SEP-2001
LOCUS
DEFINITION Sequence 58 from Patent WO0159102.
ACCESSION AX214245
VERSION AX214245.1 GI:15524322
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE
1 (bases 1 to 28)
AUTHORS Breaker,R. and Emilsson,G.
TITLE Nucleozymes with endonuclease activity
JOURNAL Patent: WO 0159102-A 58 16-AUG-2001;
RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)
FEATURES
source
1.28
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Nucleic Acid"
misc_feature
1.6
/note="2'-O-Methyl"
misc_feature
12
/note="2'-O-Methyl"
misc_feature
21..27
/note="2'-O-Methyl"
misc_feature
28
/note="n stands for inverted deoxybasic derivative"
BASE COUNT
11 a 3 c 9 g 4 t 1 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 15; DB 6; Length 28;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaagu 15
|||||:|||||:
Db 7 AGATACGTGAGAT 21

RESULT 11
AX214246 28 bp mRNA linear PAT 06-SEP-2001
LOCUS
DEFINITION Sequence 59 from Patent WO0159102.
ACCESSION AX214246
VERSION AX214246.1 GI:15524323
KEYWORDS
SOURCE synthetic construct.

ORGANISM
synthetic construct
artificial sequence.
REFERENCE
1 (bases 1 to 28)
AUTHORS Breaker,R. and Emilsson,G.
TITLE Nucleozymes with endonuclease activity
JOURNAL Patent: WO 0159102-A 59 16-AUG-2001;
RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)
FEATURES
source
1.28
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Nucleic Acid"
misc_feature
1.7
/note="2'-O-Methyl"
misc_feature
21..27
/note="2'-O-Methyl"
misc_feature
28
/note="n stands for inverted deoxybasic derivative"
BASE COUNT
11 a 3 c 9 g 4 t 1 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 15; DB 6; Length 28;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaagu 15
|||||:|||||:
Db 7 AGATACGTGAGAT 21

RESULT 12
AX214247 28 bp mRNA linear PAT 06-SEP-2001
LOCUS
DEFINITION Sequence 60 from Patent WO0159102.
ACCESSION AX214247
VERSION AX214247.1 GI:15524324
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE
1 (bases 1 to 28)
AUTHORS Breaker,R. and Emilsson,G.
TITLE Nucleozymes with endonuclease activity
JOURNAL Patent: WO 0159102-A 60 16-AUG-2001;
RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)
FEATURES
source
1.28
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Nucleic Acid"
misc_feature
1.6
/note="2'-O-Methyl"
misc_feature
21..27
/note="2'-O-Methyl"
misc_feature
28
/note="n stands for inverted deoxybasic derivative"
BASE COUNT
11 a 3 c 9 g 4 t 1 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 15; DB 6; Length 28;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaagu 15
|||||:|||||:
Db 7 AGATACGTGAGAT 21

RESULT 13
AX214248 28 bp mRNA linear PAT 06-SEP-2001
LOCUS

DEFINITION Sequence 61 from Patent WO0159102.
ACCESSION AX214248
VERSION AX214248.1 GI:15524325
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 28)
AUTHORS Breaker/R. and Emilsson,G.
TITLE Nucleozymes with endonuclease activity
JOURNAL Patent: WO 0159102-A 61 16-AUG-2001;
RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)
FEATURES
source
1..28
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Nucleic Acid"
misc_feature
1..6
/note="2'-O-Methyl"
misc_feature
18
/note="2'-O-Methyl"
misc_feature
21..27
/note="2'-O-Methyl"
misc_feature
28
/note="n stands for inverted deoxybasic derivative"
BASE COUNT 11 a 3 c 9 g 4 t 1 others
ORIGIN

Query Match 100.0%; Score 15; DB 6; length 28;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaagau 15
|||||:|||||:
Db 7 AGATTAACGTGAAGAT 21

RESULT 14
AX214249
LOCUS AX214249 28 bp mRNA linear PAT 06-SEP-2001
DEFINITION Sequence 62 from Patent WO0159102.
ACCESSION AX214249
VERSION AX214249.1 GI:15524326
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 28)
AUTHORS Breaker/R. and Emilsson,G.
TITLE Nucleozymes with endonuclease activity
JOURNAL Patent: WO 0159102-A 62 16-AUG-2001;
RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)
FEATURES
source
1..28
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Nucleic Acid"
misc_feature
1..6
/note="2'-O-Methyl"
misc_feature
17
/note="2'-O-Methyl"
misc_feature
21..27
/note="2'-O-Methyl"
misc_feature
28
/note="n stands for inverted deoxybasic derivative"
BASE COUNT 11 a 3 c 9 g 4 t 1 others
ORIGIN

Query Match 100.0%; Score 15; DB 6; length 28;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaagau 15
|||||:|||||:
Db 7 AGATTAACGTGAAGAT 21

RESULT 15
AX214250
LOCUS AX214250 28 bp mRNA linear PAT 06-SEP-2001
DEFINITION Sequence 63 from Patent WO0159102.
ACCESSION AX214250
VERSION AX214250.1 GI:15524327
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 28)
AUTHORS Breaker/R. and Emilsson,G.
TITLE Nucleozymes with endonuclease activity
JOURNAL Patent: WO 0159102-A 63 16-AUG-2001;
RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)
FEATURES
source
1..28
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Nucleic Acid"
misc_feature
1..6
/note="2'-O-Methyl"
misc_feature
11
/note="2'-O-Methyl"
misc_feature
16
/note="2'-O-Methyl"
misc_feature
21..27
/note="2'-O-Methyl"
misc_feature
28
/note="n stands for inverted deoxybasic derivative"
BASE COUNT 11 a 3 c 9 g 4 t 1 others
ORIGIN

Query Match 100.0%; Score 15; DB 6; length 28;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 agauaacgugaagau 15
|||||:|||||:
Db 7 AGATTAACGTGAAGAT 21

Search completed: September 9, 2002, 01:46:37
Job time: 9797 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 9, 2002, 01:14:06 ; Search time 57.83 Seconds

(without alignments)
76,455 Million, cell updates/sec

Title: US-09-780-929-98

Perfect score: 18

Sequence: 1 aaagcccaacgagcgca 18

Scoring table: OLIGO.MUC

Gapop 60.0 , Gapext 60.0

Searched: 38353 seqs, 122816752 residues

Word size: 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:*

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	14	77.8	92	US-09-070-842A-3	Sequence 3, Appl
C 2	14	77.8	249	US-09-070-842A-10	Sequence 10, Appl
C 3	14	77.8	372	US-09-070-842A-7	Sequence 7, Appl
C 4	14	77.8	375	US-09-070-842A-11	Sequence 11, Appl
5	13	72.2	1190	US-08-861-774E-89	Sequence 89, Appl
6	12	66.7	3536	US-09-418-640-3	Sequence 3, Appl
7	12	66.7	3597	US-09-199-637A-404	Sequence 404, App
8	12	66.7	3720	US-08-074-967-1	Sequence 1, Appl
9	12	66.7	3720	US-08-553-541B-1	Sequence 1, Appl
10	12	66.7	3720	US-09-268-202-1	Sequence 1, Appl
11	12	66.7	3720	PCR-US94-06669-1	Sequence 1, Appl
12	12	66.7	37948	US-09-251-645-11	Sequence 11, Appl
13	12	66.7	42235	US-09-199-637A-1	Sequence 1, Appl
C 14	11	61.1	170	US-08-553-619B-19	Sequence 19, Appl
C 15	11	61.1	375	US-08-793-107-7	Sequence 7, Appl
C 16	11	61.1	375	US-08-793-660B-11	Sequence 11, Appl
C 17	11	61.1	375	US-08-793-110-7	Sequence 11, Appl
C 18	11	61.1	402	US-09-199-637A-380	Sequence 380, App
C 19	11	61.1	534	US-09-199-637A-376	Sequence 376, App
20	11	61.1	540	US-09-199-637A-378	Sequence 378, App
21	11	61.1	690	US-08-926-842B-15	Sequence 15, Appl
22	11	61.1	690	US-08-926-842B-37	Sequence 37, Appl
23	11	61.1	740	US-08-771-098-1	Sequence 1, Appl
24	11	61.1	740	US-09-022-875-3	Sequence 3, Appl
25	11	61.1	740	US-09-354-040-1	Sequence 1, Appl
26	11	61.1	819	US-08-221-750A-8	Sequence 8, Appl
C 27	11	61.1	1005	US-08-482-385A-2	Sequence 2, Appl

28	11	61.1	1177	4	US-08-861-774E-23	Sequence 23, Appl
29	11	61.1	1320	1	US-08-419-414-1	Sequence 1, Appl
30	11	61.1	1383	1	US-08-484-044-1	Sequence 1, Appl
C 31	11	61.1	1435	1	US-07-968-971A-2	Sequence 2, Appl
C 32	11	61.1	1435	1	US-07-824-247-43	Sequence 43, Appl
C 33	11	61.1	1435	1	US-08-142-473A-3	Sequence 3, Appl
C 34	11	61.1	1435	1	US-08-464-923B-10	Sequence 10, Appl
C 35	11	61.1	1435	1	US-08-469-203A-3	Sequence 3, Appl
C 36	11	61.1	1435	1	US-08-470-204A-43	Sequence 43, Appl
C 37	11	61.1	1435	3	US-08-470-204A-43	Sequence 71, Appl
C 38	11	61.1	1448	4	US-08-858-207A-71	Sequence 1, Appl
39	11	61.1	1542	2	US-08-865-311-1	Sequence 1, Appl
40	11	61.1	1641	4	US-09-178-252-9	Sequence 9, Appl
41	11	61.1	1696	2	US-08-553-619B-4	Sequence 4, Appl
C 42	11	61.1	1796	1	US-08-540-077-2	Sequence 2, Appl
C 43	11	61.1	1899	4	US-08-965-762-10	Sequence 10, Appl
C 44	11	61.1	2060	3	US-09-039-555B-19	Sequence 19, Appl
45	11	61.1	2027	1	US-08-150-203A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-070-842A-3/c
Sequence 3, Application US/09070842A
Patent No. 6127171
GENERAL INFORMATION:
APPLICANT: Sillaty, N. Steve
APPLICANT: Lebel, Suzanne
TITLE OF INVENTION: Modified lacZ Coding Sequences
TITLE OF INVENTION: And Uses Thereof
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
STREET: 1800 One Met Plaza
CITY: Buffalo
STATE: New York
COUNTRY: United States
ZIP: 14203-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070, 842A
FILING DATE: 1 May 1998
ATTORNEY/AGENT INFORMATION:
NAME: Kadle, Ranjana
REGISTRATION NUMBER: 40,041
REFERENCE/DOCKET NUMBER: 24945.0001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 3 :
SEQUENCE CHARACTERISTICS:
LENGTH: 92 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single-stranded
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
US-09-070-842A-3

Query Match 77.8%; Score 14; DB 3; Length 92;
Best local Similarity 78.6%; Pred. No. 1.6;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 ggcacacgagcgcg 17
||||:|||||
Db 55 ggcctatcgctgcg 42

RESULT 2
US-09-070-842A-10/c
; Sequence 10, Application US/09070842A
; Patent No. 6127171
; GENERAL INFORMATION:
; APPLICANT: Siliaty, N. Steve
; APPLICANT: Lebel, Suzanne
; TITLE OF INVENTION: Modified lacZ' Coding Sequences
; TITLE OF INVENTION: And Uses Thereof
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,842A
; FILING DATE: 1 May 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Kadle, Ranjana
; REGISTRATION NUMBER: 40,041
; REFERENCE/DOCKET NUMBER: 24945.0001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 856-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 10 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 249 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
; HYPOTHEICAL: NO
; FEATURE: relevant portion of circular molecule listed
; US-09-070-842A-10

Query Match 77.8%; Score 14; DB 3; Length 249;
Best Local Similarity 78.6%; Pred. No. 1.6;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 ggcuaucgugcg 17
||||:||||:
DB 188 ggcctatcgctcg 175

RESULT 3
US-09-070-842A-7/c
; Sequence 7, Application US/09070842A
; Patent No. 6127171
; GENERAL INFORMATION:
; APPLICANT: Siliaty, N. Steve
; APPLICANT: Lebel, Suzanne
; TITLE OF INVENTION: Modified lacZ' Coding Sequences
; TITLE OF INVENTION: And Uses Thereof
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/ Microsoft Windows
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,842A
FILING DATE: 1 May 1998
ATTORNEY/AGENT INFORMATION:
NAME: Kadle, Ranjana
REGISTRATION NUMBER: 40,041
REFERENCE/DOCKET NUMBER: 24945.0001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 856-4000
TELEFAX: (716) 849-0349
INFORMATION FOR SEQ ID NO: 7 :
SEQUENCE CHARACTERISTICS:
LENGTH: 372 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double-stranded
TOPOLOGY: circular
MOLECULE TYPE: DNA
HYPOTHEICAL: NO
FEATURE: relevant portion of circular molecule listed
; US-09-070-842A-7

Query Match 77.8%; Score 14; DB 3; Length 372;
Best Local Similarity 78.6%; Pred. No. 1.6;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 ggcuaucgugcg 17
||||:||||:
DB 233 ggcctatcgctcg 220

RESULT 4
US-09-070-842A-11/c
; Sequence 11, Application US/09070842A
; Patent No. 6127171
; GENERAL INFORMATION:
; APPLICANT: Siliaty, N. Steve
; APPLICANT: Lebel, Suzanne
; TITLE OF INVENTION: Modified lacZ' Coding Sequences
; TITLE OF INVENTION: And Uses Thereof
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,842A
; FILING DATE: 1 May 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Kadle, Ranjana
; REGISTRATION NUMBER: 40,041
; REFERENCE/DOCKET NUMBER: 24945.0001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 856-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 11 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: circular

MOLECULE TYPE: DNA
HYPOTHETICAL: NO
FEATURE: relevant portion of circular molecule listed
US-09-070-842A-11

Query Match
Best Local Similarity 77.8%; Score 14; DB 3; Length 375;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 4 ggcgaucgucgucg 17
||||:||||:||||
DB 233 GGCCTATCGCTCGC 220

RESULT 5
US-08-861-774E-89
Sequence 89, Application US/08861774E
Patent No. 6297007
GENERAL INFORMATION:
APPLICANT: Waters, Barbara
APPLICANT: Miao, Vivian
APPLICANT: Ho, Yap
APPLICANT: Tong, Seow
TITLE OF INVENTION: METHOD FOR ISOLATION OF BIOSYNTHESIS GENES FOR
TITLE OF INVENTION: BIOACTIVE MOLECULES
FILE REFERENCE: 9993-006
CURRENT APPLICATION NUMBER: US/08/861,774E
CURRENT FILING DATE: 1997-05-22
NUMBER OF SEQ ID NOS: 94
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 89
LENGTH: 1190
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Clone ps24
US-08-861-774E-89

Query Match
Best Local Similarity 72.2%; Score 13; DB 4; Length 1190;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 aaugccuacg 13
||:||||:||||
DB 214 aatgcctatcg 226

RESULT 6
US-09-418-640-3
Sequence 3, Application US/09418640
Patent No. 6140125
GENERAL INFORMATION:
APPLICANT: Jennifer K. Taylor
APPLICANT: Lex M. Cowser
TITLE OF INVENTION: ANTISENSE MODULATION OF BCL-6 EXPRESSION
FILE REFERENCE: RTS-0102
CURRENT APPLICATION NUMBER: US/09/418,640
CURRENT FILING DATE: 1999-10-15
NUMBER OF SEQ ID NOS: 89
SEQ ID NO 3
LENGTH: 3536
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (328)..(2448)
US-09-418-640-3

Query Match
Best Local Similarity 66.7%; Score 12; DB 3; Length 3536;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
OY 2 aaugccuacg 13
||:||||:||||
DB 790 atgcctatcg 801

RESULT 7
US-09-199-637A-404
Sequence 404, Application US/09199637A
Patent No. 6355411
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick
APPLICANT: Goodman, Howard M.
APPLICANT: Rahme, Laurence G.
APPLICANT: Mahajan-Miklos, Shalina
APPLICANT: Tan, Man-Wah
APPLICANT: Cao, Hui
APPLICANT: Drenkard, Eliana
APPLICANT: Tsongalis, John
TITLE OF INVENTION: VIRUS-ASSOCIATED NUCLEIC ACID
TITLE OF INVENTION: SEQUENCES AND USES THEREOF
FILE REFERENCE: 00786/361002
CURRENT APPLICATION NUMBER: US/09/199,637A
CURRENT FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: 60/066,517
PRIOR FILING DATE: 1997-11-25
NUMBER OF SEQ ID NOS: 437
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 404
LENGTH: 3597
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-404

Query Match
Best Local Similarity 66.7%; Score 12; DB 4; Length 3597;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 aaugccuacg 12
||:||||:||||
DB 2614 aatgcctatcg 2625

RESULT 8
US-08-074-967-1
Sequence 1, Application US/08074967
Patent No. 5641672
GENERAL INFORMATION:
APPLICANT: Dalla-Favera, Riccardo
APPLICANT: Chaganti, R.S.K.
TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/074,967
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.

REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/43771
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3720 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 328..2445
US-08-074-967-1

Query Match 66.7%; Score 12; DB 1; Length 3720;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 2 auggccuacgg 13
|:||||:|:||||
Db 790 ATGGCCTATCGG 801

RESULT 9
US-08-553-541B-1
Sequence 1, Application US/08553541B
Patent No. 5882858
GENERAL INFORMATION:
APPLICANT: Dalla-Favera, Riccardo
APPLICANT: Chaganti, Raju S.K.
TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
TITLE OF INVENTION: bcl-6
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/553,541B
FILING DATE: May 28, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/43771-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3720 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 328..2445
US-08-553-541B-1

Query Match 66.7%; Score 12; DB 2; Length 3720;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 2 auggccuacgg 13
|:||||:|:||||
Db 790 ATGGCCTATCGG 801

RESULT 10
US-09-268-202-1
Sequence 1, Application US/09268202
Patent No. 6174997
GENERAL INFORMATION:
APPLICANT: Dalla-Favera, Riccardo
APPLICANT: Chaganti, Raju S.K.
TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
TITLE OF INVENTION: bcl-6
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/268,202
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/43771-A-PCT-US-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3720 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 328..2445
US-09-268-202-1

Query Match 66.7%; Score 12; DB 4; Length 3720;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 2 auggccuacgg 13
|:||||:|:||||
Db 790 ATGGCCTATCGG 801

RESULT 11
PCT-US94-06669-1
Sequence 1, Application PC/TUS9406669
GENERAL INFORMATION:
APPLICANT: Dalla-Favera, Riccardo
APPLICANT: Chaganti, R.S.K.

;;
TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
TITLE OF INVENTION: bcl-6
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/06669
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/074,967
FILING DATE: 09-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 43771-A-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3720 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 328..2445
PCT-US94-06669-1

Query Match 66.7%; Score 12; DB 5; Length 3720;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 aaugccuacug 13
1:||||:|:|

DB 790 atggcctatcg 801

RESULT 12
US-09-251-645-11
; Sequence 11, Application US/09251645
; Patent No. 6281413
; GENERAL INFORMATION:
; APPLICANT: Kramer, Vance C.
; APPLICANT: Morgan, Michael K.
; APPLICANT: Anderson, Arne R.
; APPLICANT: Hart, Hope
; APPLICANT: Warren, Gregory W.
; APPLICANT: Dunn, Martha
; APPLICANT: Chen, Jeng S.
; TITLE OF INVENTION: NOVEL INSECTICIDAL TOXINS FROM PHOTORHABDUS LUMINESCENS
; TITLE OF INVENTION: AND NUCLEIC ACID SEQUENCES CODING THEREFOR
; FILE REFERENCE: CGC1963/A
; CURRENT APPLICATION NUMBER: US/09/251,645
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 37948

;;
TYPE: DNA
ORGANISM: Photorhabdus luminescens
FEATURE:
NAME/KEY: CDS
LOCATION: (1517)..(18035)
OTHER INFORMATION: orf5
FEATURE:
NAME/KEY: CDS
LOCATION: (23768)..(31336)
OTHER INFORMATION: hph2
FEATURE:
NAME/KEY: CDS
LOCATION: (31393)..(35838)
OTHER INFORMATION: orf2
US-09-251-645-11

Query Match 66.7%; Score 12; DB 4; Length 37948;
Best Local Similarity 75.0%; Pred. No. 27;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaugccuacug 12
1:||||:|:|

DB 39382 aatgcctatcg 39393

RESULT 14
US-08-553-619B-19/c
; Sequence 19, Application US/08553619B
; Patent No. 5919705
; GENERAL INFORMATION:
; APPLICANT: Tsongalis, John
; APPLICANT: Drenkard, Eliana
; APPLICANT: Cao, Hui
; APPLICANT: Tan, Man-Wah
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Gao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 42235
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(42235)
; OTHER INFORMATION: N is any nucleic acid.
US-09-199-637A-1

Query Match 66.7%; Score 12; DB 4; Length 42235;
Best Local Similarity 75.0%; Pred. No. 27;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaugccuacug 12
1:||||:|:|

DB 39382 aatgcctatcg 39393

RESULT 14
US-08-553-619B-19/c
; Sequence 19, Application US/08553619B
; Patent No. 5919705
; GENERAL INFORMATION:

APPLICANT: Dehaan, Petrus T.
TITLE OF INVENTION: Virus Resistant Plants
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5919705artis Crop Protection
STREET: 975 California Avenue
City: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/553,619B
FILING DATE: December 1, 1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Marcus-Wyner, Lynn
REGISTRATION NUMBER: 34,869
REFERENCE/DOCKET NUMBER: 137-1082/PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/354-3588
TELEFAX: 415/857-1125
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: oligonucleotide
US-08-553-619B-19

Query Match 61.1%; Score 11; DB 2; Length 170;
Best Local Similarity 72.7%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 aaugccuanc 11
1:|||||:1
DB 152 AATGGCTATC 142

RESULT 15
US-08-793-107-7/c
Sequence 7, Application US/08793107
Patent No. 5830880
GENERAL INFORMATION:
APPLICANT: SEDLACEK, Hans-Harald
APPLICANT: BOSSLET, Klaus
APPLICANT: MUELLER, Rolf
TITLE OF INVENTION: GENE THERAPY OF TUMORS WITH AN
TITLE OF INVENTION: ENDOTHELIAL CELL-SPECIFIC, CELL CYCLE-DEPENDENT ACTIVE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
STATE: D.C.
CITY: Washington
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,107

FILING DATE: 18-APR-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: NO PCT/EP95/03370
FILING DATE: 25-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9506466.3
FILING DATE: 29-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9417366.3
FILING DATE: 26-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: GRANADOS, Patricia D.
REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 18748/333
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-793-107-7

Query Match 61.1%; Score 11; DB 2; Length 375;
Best Local Similarity 72.7%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 auggccuanc 12
1:|||||:1
DB 224 AATGGCTATC 214

Search completed: September 9, 2002, 01:14:08
Job time: 10293 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: September 9, 2002, 01:12:49 ; Search time 2186.81 seconds

(without alignments)
111.096 Million cell updates/sec

Title: US-09-780-929-98

Perfect score: 18

Sequence: 1 aaugcccaugcgagcga 18

Scoring table: OLIGO_NUC

Searched: 13736207 segs, 6748477542 residues

Word size: 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: EST:*

1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estrov:*
6: em_estcpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	83.3	556	12	A2147327 SP_0019_A
2	15	83.3	664	12	BH342700 CH230-890
3	15	83.3	966	10	BF584504
4	14	77.8	422	9	AW426165
5	14	77.8	456	12	AQ306902 HS_3080_A
6	14	77.8	497	10	BJ001054
7	14	77.8	509	10	BT782695
8	14	77.8	537	10	BE268497
9	14	77.8	585	10	BE268497
10	14	77.8	599	9	BJ040691
11	14	77.8	650	10	BE368419
12	14	77.8	685	10	BJ015614
13	14	77.8	728	10	BJ055067
14	14	77.8	773	10	BJ022373
15	14	77.8	802	12	CNS0725Y
16	14	77.8	817	12	A2195692
17	14	77.8	847	12	A2195694

C 18	14	77.8	906	12	CNS0203R
C 19	14	77.8	915	10	BG671636
C 20	14	77.8	927	12	CNS0280H
C 21	14	77.8	965	12	CNS02PGL
C 22	14	77.8	980	10	BG284272
C 23	14	77.8	998	12	CNS0601L
C 24	14	77.8	1009	10	BG562941
C 25	14	77.8	1016	10	BG542974
C 26	14	77.8	1095	10	BI525445
C 27	14	77.8	1297	10	BE420957
C 28	13	72.2	172	9	AI895547
C 29	13	72.2	247	9	AV341677
C 30	13	72.2	263	9	BB579496
C 31	13	72.2	265	9	AI641968
C 32	13	72.2	275	9	AV627152
C 33	13	72.2	331	10	BF377114
C 34	13	72.2	369	9	AA763830
C 35	13	72.2	374	9	AA075662
C 36	13	72.2	401	9	AA062098
C 37	13	72.2	408	9	AA451482
C 38	13	72.2	419	10	BF069364
C 39	13	72.2	427	12	AQ248758
C 40	13	72.2	428	10	BM357310
C 41	13	72.2	428	12	AQ771282
C 42	13	72.2	431	9	AA500416
C 43	13	72.2	432	10	BM481429
C 44	13	72.2	433	10	BF755781
C 45	13	72.2	434	9	BB863683

ALIGNMENTS

RESULT 1
A2147327
LOCUS SP_0019_A1.D07.T7 Strongylocentrotus purpuratus, purple sea urchin,
DEFINITION sperm genomic BAC library Strongylocentrotus purpuratus genomic
clone plate-19 Col-13 Row-G, DNA sequence.

ACCESSION A2147327 556 bp DNA linear GSS 28-AUG-2000
VERSION A2147327.1 GI:8299228
KEYWORDS GSS.
SOURCE Strongylocentrotus purpuratus.
ORGANISM Strongylocentrotus purpuratus.
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinozoa; Echinodermata; Echinozoa; Echinodermata; Echinozoa;
Strongylocentrotidae; Strongylocentrotus.
1 (bases 1 to 556)
Cameron, R.A., Mahairas, G., Rast, J.P., Martinez, P., Biondi, T.R.,
Swartzell, S., Wallace, J.C., Poustka, A.J., Livingston, B.T., May
G.A., Eitensohn, C.A., Lehman, H., Britten, R.J., Davidson, E.H. and
Hood, L.

TITLE A sea urchin genome project: Sequence scan, virtual map, and
additional resources
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 97 (17), 9514-9518 (2000)
MEDLINE 20402566

COMMENT Contact: Cameron, RA, Davidson, EH, Hood, L
Division of Biology 156-29
California Institute of Technology
Pasadena California 91125, USA
Tel: (626) 395-8421
Fax: (626) 793-3047
Email: acameron@caltech.edu
Plate: 19 row: G column: 13
Seg primer: T7
Class: BAC ends
High quality sequence stop: 556.

FEATURES

source Location/Qualifiers
1..556
/organism="Strongylocentrotus purpuratus"
/db_xref="taxon:7668"
/clone="Plate-19 Col-13 Row-G"
/clone_1lb="Strongylocentrotus purpuratus, purple sea

urchn, sperm genomic BAC library"
/note="Organ: sperm; Vector: BAC3.6; BAC Clones in E-Coli
DH10B"

BASE COUNT 147 a 105 c 113 g 188 t 3 others
ORIGIN

Query Match 83.3%; Score 15; DB 12; Length 556;
Best Local Similarity 73.3%; Pred. No. 25;
Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 aaugccuacugugc 15
11:1111:1111:1
Db 152 AATGCCATCGCTGC 166

RESULT 2
BH342700 664 bp DNA linear GSS 03-DEC-2001
LOCUS CH230-89016 TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
DEFINITION CH230-89016, DNA sequence.
ACCESSION BH342700
VERSION BH342700.1 GI:17273434
KEYWORDS GSS.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 664)
Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K., Shvartsbeyn
, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D., Riggs, F., de
Jong, P. and Fraser, C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)
Other GSSs: CH230-89016.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhaoc@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@tigr.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/orering/information.htm). BAC end
page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
Plate: 89 row: 0 column: 16
Seq primer: SP6
Class: BAC ends.

FEATURES
Source location/Qualifiers
1..664
/organism="Rattus norvegicus"
/strain="BN/SSNHsd/MCM"
/db_xref="taxon:10116"
/clone="CH230-89016"
/clone_lib="CHORI-230 Segment 1"
/sex="Female"
/cell_type="Brain"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SSNHsd/MCM) BAC library produced by
Pieter de Jong"

BASE COUNT 190 a 157 c 146 g 171 t
ORIGIN

Query Match 83.3%; Score 15; DB 12; Length 664;
Best Local Similarity 73.3%; Pred. No. 26;
Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 aaugccuacugugc 16

111111:111111:11
Db 408 ATGCCATCGCTGC 422

RESULT 3
BF584504 966 bp mRNA linear EST 12-DEC-2000
LOCUS 602098202F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218168 5',
DEFINITION mRNA sequence.
ACCESSION BF584504
VERSION BF584504.1 GI:11658222
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 966)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapds@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9798 row: c column: 01
High quality sequence stop: 656.

FEATURES
Source location/Qualifiers
1..966
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4218168"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP library."

BASE COUNT 280 a 281 c 226 g 178 t 1 others
ORIGIN

Query Match 83.3%; Score 15; DB 10; Length 966;
Best Local Similarity 80.0%; Pred. No. 28;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 4 ggcuccuacugugcga 18
111111:111111:1111
Db 911 GGCCATCGCTGCGA 925

RESULT 4
AA426165 422 bp mRNA linear EST 25-APR-2001
LOCUS 59757 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION AA426165
ACCESSION AA426165
VERSION AA426165.1 GI:6954112
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
1 (bases 1 to 422)
Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett,
G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,
Perte, G., Holt, L., Karameycheva, S., Liang, F., Quackenbush, J. and

TITLE
JOURNAL
MEDLINE
COMMENT

Keefe, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smithemail.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross-match with the -minscore 20 and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACGCTATGACCAT
BACKWARD: GTTTCCACGTACGACG
Plate: 27 row: 3 column: 16
Seq primer: ATTGAGTACGACATATG.
Location/Qualifiers
1. 422
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 4BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site 1: XbaI; Site 2: XhoI; library made from pooled tissue from day 20 and day 40 embryos."
BASE COUNT
66 a 120 c 137 g 99 t
ORIGIN

Query Match
Best Local Similarity 77.8%; Score 14; DB 9; Length 422;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ugccuacugcugc 16
:||||:||||:|
Db 246 TGCCCTATCGGTGC 259

RESULT 5
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AQ306902
HS_3080.AL.D12_MR_CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3080 Col=23 Row=G, DNA sequence.
AQ306902.1 GI:4026688
GSS.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 456)
Mahairs, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallaceu.washington.edu
Sequence Tagged Connector
Plate: 3080 row: G column: 23
Class: BAC ends
High quality sequence stop: 456.

FEATURES
SOURCE
Location/Qualifiers
1. 456
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=3080 Col=23 Row=G"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

BASE COUNT
130 a 87 c 104 g 131 t 4 others
ORIGIN

Query Match
Best Local Similarity 77.8%; Score 14; DB 12; Length 456;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ugccuacugcugc 16
:||||:||||:|
Db 266 TGCCCTATCGGTGC 279

RESULT 6
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BJ001054
BJ001054 MF01SSA cDNA Oryzias latipes cDNA clone MF01SSA024E06 5', mRNA sequence.
BJ001054
BJ001054.1 GI:17359024
EST.
Japanese medaka.
Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorphi; Atherinomorphi; Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
1 (bases 1 to 497)
Kohara, Y., Shin-I, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.
Medaka EST Project in Takeda's lab
unpublished (2001)
Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
1. 497
/organism="Oryzias latipes"
/strain="Hd-rf"
/db_xref="taxon:8090"
/clone_lib="MF01SSA024E06"
/clone="MF01SSA cDNA"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="segmentation stage 20 - 25"

BASE COUNT
157 a 114 c 76 g 149 t 1 others
ORIGIN

Query Match
Best Local Similarity 77.8%; Score 14; DB 10; Length 497;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ugccuacugcugc 16
:||||:||||:|
Db 154 TGCCCTATCGGTGC 141

RESULT 7
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BJ782695
509 bp mRNA linear EST 26-SEP-2001

DEFINITION kh41b05.v1 Ascaris suum male gut PAMP1 v2 Chiapelli McCarter
Ascaris suum cDNA 5', mRNA sequence.
ACCESSION BE269497
VERSION BE269497.1 GI:15785587
KEYWORDS EST.
SOURCE pig roundworm.
ORGANISM Ascaris suum
Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascaridoidea
; Ascarididae; Ascaris.
REFERENCE 1 (bases 1 to 509)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.,
Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.,
Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagaris,I., R.,
Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,
Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
Wilson,R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
COMMENT Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
The library was constructed by Brandi Chiapelli and Dr. James
McCarter at Washington University, St. Louis. The cDNA was made by
using Dynabead oligo-dT priming (Dyna1). PCR based library using a
modified protocol from the SMART PCR cDNA Synthesis Kit from
Clontech. Directionally cloned into the UDG sites of PAMP1.
Dissected nematode tissues were provided by Dr. Alan Scott
(ascott@hsp.edu) of the School of Public Hygiene and Public Health
at John Hopkins University in Baltimore, MD.
Seq primer: -40RP from Gibco
High quality sequence stop: 395.
Location/Qualifiers
FEATURES
Source
1..509
/organism="Ascaris suum"
/db_xref="taxon:6253"
/clone_id="Ascaris suum male gut PAMP1 v2 Chiapelli
McCarter"
/sex="Male"
/tissue_type="Intestine"
/dev_stage="Adult"
/lab_host="DH10B"
/note="Vector: PAMP1 (Gibco); Site_1: NotI; Site_2: SalI;
The library was constructed by Brandi Chiapelli and Dr.
James McCarter at Washington University, St. Louis. The
cDNA was made by using Dynabead oligo-dT priming (Dyna1).
PCR based library using a modified protocol from the SMART
PCR cDNA Synthesis Kit from Clontech. Directionally cloned
into the UDG sites of PAMP1. Dissected nematode tissues
were provided by Dr. Alan Scott (ascott@hsp.edu) of the
School of Public Hygiene and Public Health at John Hopkins
University in Baltimore, MD."
BASE COUNT 169 a 117 c 75 g 148 t
ORIGIN
Query Match 77.8%; Score 14; DB 10; Length 509;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Oy 3 uggcuaucgugc 16
:||||:|||||
Db 476 TGGCCTATCGTGC 489
RESULT 8
BE269497 537 bp mRNA linear EST 13-JUL-2000
LOCUS
DEFINITION 60118465F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3542550 5',

ACCESSION BE269497
VERSION BE269497.1 GI:9143118
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 537)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps+email.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL).
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM234 row: m column: 23
High quality sequence stop: 321.
Location/Qualifiers
FEATURES
Source
1..537
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3542550"
/clone_id="NIH_MGC_8"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pOT7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using 2AP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 119 a 120 c 160 g 136 t 2 others
ORIGIN
Query Match 77.8%; Score 14; DB 10; Length 537;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Oy 1 aaugcuaucggu 14
:||||:|||||
Db 509 AATGGCCTATCGGT 522
RESULT 9
BU040691 585 bp mRNA linear EST 06-DEC-2001
LOCUS
DEFINITION BU040691 NIBB Mochli normalized Xenopus neurola library Xenopus
laevis cDNA clone X1047e11 5', mRNA sequence.
ACCESSION BU040691
VERSION BU040691.1 GI:17388082
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 585)
AUTHORS Kitayama,A., Terasaka,C., Mochli,M., Ueno,N., Shin-I,T. and Kohara
Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
CONTACT: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers

FEATURES

source

1.585

/organism="Xenopus laevis"
/db_xref="taxon:8355"

/clone_lib="NIBB Mochii normalized Xenopus neurula
library"

/tissue_type="whole embryo"

/dev_stage="stage 15"

BASE COUNT

180 a 114 c 135 g 153 t 3 others

ORIGIN

Query Match

Best Local Similarity 77.8%; Score 14; DB 10; Length 585;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 5 ggcuaucgugcg 18

Db 254 GCCTATCGGTGCGA 267

RESULT 10

AL643798

LOCUS AL643798 XGC-gastrula silurana tropicalis cDNA clone LIG3h2 5',
DEFINITION mRNA sequence.

ACCESSION

AL643798.1 GI:16795923

VERSION

KEYWORDS

SOURCE

ORGANISM

Western clawed frog.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;

Xenopodinae; Silurana.

1 (bases 1 to 599)
Huckle, E., Taylor, R., Ashurst, J.L., Zorn, A.M. and Rogers, J.
Sanger Xenopus tropicalis EST project 2001 (10_2001)
Unpublished (2001)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Huckle E

Sanger Centre

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS_SEQUENCE_ID: LIG3h2.sp6

Sequencing primer: SP6

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Aaron M. Zorn.

Location/Qualifiers

1.599

/organism="Silurana tropicalis"

/db_xref="taxon:8364"

/clone="LIG3h2"

/clone_lib="XGC-gastrula"

/dev_stage="gastrula (stages 10-5-13 mixed)"

/lab_host="Escherichia coli XL1-blue"

/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA

was oligo dt primed from 5ug of poly A+ RNA from stages

10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated

into pCS107 with EcoRI at the 5' end and NotI at the 3'

end."

BASE COUNT 110 a 157 c 201 g 130 t 1 others

ORIGIN

Query Match 77.8%; Score 14; DB 9; Length 599;

Best Local Similarity 71.4%; Pred. No. 1.1e+02;

Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 3 ugccuaucgugcg 16

Db 320 TGGCTATCGGTGCG 333

RESULT 11

BE368419/c

LOCUS BE368419 650 bp mRNA linear EST 21-JUL-2000

DEFINITION 601220492F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3589319 5',
mRNA sequence.

ACCESSION

BE368419

VERSION

BE368419.1 GI:9313782

KEYWORDS

EST.

house mouse.

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 650)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LLM8755 row: h column: 24

High quality sequence start: 6

High quality sequence stop: 477.

Location/Qualifiers

1.650

/organism="Mus musculus"

/strain="C57BL/6J (f1)"

/db_xref="taxon:10090"

/clone="IMAGE:3589319"

/clone_lib="NCI_CGAP_Lu29"

/tissue_type="spontaneous tumor, metastatic to mammary.

/lab_host="DH10B"

/note="Organ: Lung; Vector: pCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally; Primer: Oligo dt.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

BASE COUNT 124 a 168 c 190 g 168 t

ORIGIN

Query Match 77.8%; Score 14; DB 10; Length 650;

Best Local Similarity 78.6%; Pred. No. 1.1e+02;

Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 4 ggcuaucgugcg 17

Db 43 GGCCTATCGGTGCG 30

RESULT 12

BJ015614

LOCUS BJ015614 685 bp mRNA linear EST 07-DEC-2001

DEFINITION BJ015614 MFO15SA cDNA Oryzias latipes cDNA clone MFO15SA024E06 3',
mRNA sequence.

ACCESSION BJ015614

VERSION BJ015614.1 GI:17410165

KEYWORDS

SOURCE

ORGANISM

Oryzias latipes

Japanese medaka.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Alburnomorpha;

Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

REFERENCE 1 (bases 1 to 685)
 AUTHORS Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
 TITLE Medaka EST Project in Takeda's lab
 JOURNAL Unpublished (2001)
 COMMENT Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.
 Location/Qualifiers
 1. .685
 /organism="Oryzias latipes"
 /strain="Hd-rR"
 /db_xref="taxon:8090"
 /clone_lib="MF01SSA024E06"
 /clone_lib="MF01SSA cDNA"
 /sex="mixture of female and male"
 /tissue_type="whole embryo"
 /dev_stage="segmentation stage 20 - 25"
 241 a 100 c 130 g 214 t

FEATURES

source

Query Match 77.8%; Score 14; DB 10; Length 685;
 Best Local Similarity 71.4%; Pred. No. 1.1e+02;
 Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 3 ugccuauccgugc 16
 :||||:||||:
 Db 634 TGGCCTATCGGTGC 647

RESULT 13
 BJ055067/c 728 bp mRNA linear EST 07-DEC-2001
 LOCUS
 DEFINITION BJ055067 NIBB Mochii normalized Xenopus neurula library Xenopus laevis cDNA clone XL047ell 3', mRNA sequence.
 ACCESSION BJ055067
 VERSION BJ055067.1 GI:17421319
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipelidae; Pipidae; Xenopodinae; Xenopus.

REFERENCE 1 (bases 1 to 728)
 AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara,Y.
 TITLE Expressed genes in X. laevis embryo
 JOURNAL Unpublished (2001)
 COMMENT Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.
 Location/Qualifiers
 1. .728
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="XL047ell"
 /clone_lib="NIBB Mochii normalized Xenopus neurula library"
 /tissue_type="whole embryo"
 /dev_stage="stage 15"
 206 a 147 c 135 g 239 t 1 others

FEATURES

source

Query Match 77.8%; Score 14; DB 10; Length 728;

Best Local Similarity 78.6%; Pred. No. 1.1e+02;
 Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 5 gccuauccgugcga 18
 :||||:||||:
 Db 724 GCCTATCGGTGCGA 711

RESULT 14

BJ022373

LOCUS
 DEFINITION BJ022373 MF01SSA cDNA Oryzias latipes cDNA clone MF01SSA122A06 3', mRNA sequence.
 ACCESSION BJ022373
 VERSION BJ022373.1 GI:17373772
 KEYWORDS EST.
 SOURCE Japanese medaka.
 ORGANISM
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; Belontiiformes; Adrianiichthyidae; Oryziinae; Oryzias.

REFERENCE

1 (bases 1 to 773)

AUTHORS Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.

TITLE Medaka EST Project in Takeda's lab

JOURNAL Unpublished (2001)

COMMENT Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.
 Location/Qualifiers
 1. .773
 /organism="Oryzias latipes"
 /strain="Hd-rR"
 /db_xref="taxon:8090"
 /clone="MF01SSA122A06"
 /clone_lib="MF01SSA cDNA"
 /sex="mixture of female and male"
 /tissue_type="whole embryo"
 /dev_stage="segmentation stage 20 - 25"
 267 a 112 c 146 g 248 t

FEATURES

source

Query Match 77.8%; Score 14; DB 10; Length 773;
 Best Local Similarity 71.4%; Pred. No. 1.1e+02;
 Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 3 ugccuauccgugc 16
 :||||:||||:
 Db 669 TGGCCTATCGGTGC 682

RESULT 15

CNS072SY

LOCUS
 DEFINITION CNS072SY 802 bp DNA linear GSS 07-JUL-2001
 clone BA0AB014A12 of library BA0AB from strain CLIB 210 of Kluyveromyces lactis, genomic survey sequence.
 ACCESSION AL426632
 VERSION AL426632.1 GI:12209826
 KEYWORDS GSS.
 SOURCE Kluyveromyces lactis.
 ORGANISM
 Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Kluyveromyces.

REFERENCE

1 (bases 1 to 802)

AUTHORS Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F., Duchateau-Nguyen,G., Lemaire,M., Marmeisse,R., Montrocher,R., Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.

Genomic exploration of the hemiascomycetous yeasts: 11.

Kluyveromyces lactis
 JOURNAL FEBS Lett. 487 (1), 66-70 (2000)
 MEDLINE 20584721
 REFERENCE 2 (bases 1 to 802)
 AUTHORS Soucié,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
 Bolotin-Fukuhara,M., Bon,E., Brothier,P., Casaregola,S.,
 de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
 Malpertuy,A., Neuvéglise,C., Ozler-Kalogeropoulos,O., Potier,S.,
 Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
 Winkler,P. and Weissenbach,J.
 TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
 yeast species for molecular evolution studies
 JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
 MEDLINE 20584711
 REFERENCE 3 (bases 1 to 802)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage,
 2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
 seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 COMMENT This GSS is part of a random genomic sequencing program of thirteen
 yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
 exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
 Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
 lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
 angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
 Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
 5 kb were prepared and both extremities were sequenced. See
 keywords for description of this sequence and for the sequence of
 the other extremity of this insert.
 FEATURES
 source location/Qualifiers
 1..802
 /organism="Kluyveromyces lactis"
 /strain="ClIB 210"
 /variety="lactis"
 /db_xref="taxon:28985"
 /clone="BA0AB014A12"
 /clone_1lb="BA0AB"
 <2..>800
 /note="similar to Saccharomyces cerevisiae ORF YPR160w [GPH1 : glycogen phosphorylase]"
 /evidence="not_experimental"
 misc_feature
 BASE COUNT 242 a 150 c 172 g 225 t 13 others
 ORIGIN
 Query Match 77.8%; Score 14; DB 12; Length 802;
 Best Local Similarity 71.4%; Pred. No. 1.1e+02;
 Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 1 aaagccuauaggu 14
 ||:||||:|:|:|:
 Db 631 AATGGCTATCGCT 644

Search completed: September 9, 2002, 01:12:53
 Job time: 10279 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 9, 2002, 01:51:27 ; Search time 268.97 Seconds

(without alignments)
114.899 Million cell updates/sec

Title: US-09-780-929-98

Perfect score: 18

Sequence: 1 aauggcuaucgugcgca 18

Scoring table: OLIGO-MUC

Gapex 60.0 , Gapext 60.0

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802.*

1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	18	22	AA512348
2	18	100.0	29	22	AA512381
3	14	77.8	249	20	AAV99664
4	14	77.8	249	21	AA61661
5	14	77.8	372	20	AAV99663
6	14	77.8	372	21	AA61660
7	14	77.8	375	20	AAV99665
8	14	77.8	375	21	AA61662
9	14	77.8	3839	21	AA53641

C	10	13	72.2	415	22	AA55950
C	11	13	72.2	652	21	AA609281
C	12	13	72.2	1011	22	AA667911
C	13	13	72.2	1053	22	AA671301
C	14	13	72.2	1128	22	AA621153
C	15	13	72.2	1128	22	AA530162
C	16	13	72.2	1190	20	AAV69561
C	17	13	72.2	1221	23	AA54074
C	18	13	72.2	1562	22	AAH14940
C	19	13	72.2	1968	23	AA588435
C	20	13	72.2	2667	23	ABL22419
C	21	13	72.2	6023	23	ABL22418
C	22	13	72.2	7960	10	AAV90568
C	23	13	72.2	39003	22	AA628534
C	24	13	72.2	42035	23	AA559562
C	25	13	72.2	349980	22	AAH68533
C	26	12	66.7	473	22	AAH29146
C	27	12	66.7	483	22	AAH89364
C	28	12	66.7	627	20	AAV87233
C	29	12	66.7	651	21	AA612958
C	30	12	66.7	682	23	ABL25285
C	31	12	66.7	789	20	AAZ16233
C	32	12	66.7	824	22	AAH00319
C	33	12	66.7	849	22	AAH96296
C	34	12	66.7	1005	22	AAH65832
C	35	12	66.7	1005	22	AAH52444
C	36	12	66.7	1182	22	AA668074
C	37	12	66.7	1332	9	AAH80042
C	38	12	66.7	1332	13	AAQ23035
C	39	12	66.7	1428	23	AA552466
C	40	12	66.7	1645	20	AAV82456
C	41	12	66.7	1926	23	AA580027
C	42	12	66.7	2054	21	AA679642
C	43	12	66.7	2058	21	AA675188
C	44	12	66.7	2196	22	AAH18869
C	45	12	66.7	2714	23	ABL20468

ALIGNMENTS

RESULT 1
AA512348
ID AA512348 standard; DNA; 18 BP.
AC AA512348;
XX
DT 21-NOV-2001 (first entry)
XX
DE DNA encoding deoxyribozyme #8.
XX
KW Deoxyribozyme; cytosstatic; endonuclease; RNA cleavage; DNA cleavage;
KW gene therapy; plant; fungus; bacteria; mammal; ribozyme; ss.
XX
OS Synthetic.
XX
PN WO200159102-A2.
XX
PD 16-AUG-2001.
XX
PF 08-FEB-2001; 2001WO-US04223.
XX
PR 08-FEB-2000; 2000US-0181360.
PR 31-MAR-2000; 2000US-0193646.
XX
PA (RIBO-) RIBOZYME PHARM INC.
PA (UYVA) UNIV YALE.
XX
PI Breaker R, Beigelman L, Emilsson G;
XX WPI; 2001-536526/59.
DR
XX
PT New nucleic acids with endonuclease activity, such as ribozymes and

PT nucleozymes, for modulating gene expression in a plant, mammalian,
PT bacterial or fungal cell -
PS Claim 49, Page 77, 96pp; English.
XX
XX
CC The invention relates to nucleic acid molecules with endonuclease
CC activity, which are particularly useful for cleavage of RNA or DNA.
CC The nucleic acids are used in a pharmaceutical composition and are used
CC to modulate expression of a gene in a plant, mammalian, bacterial or
CC fungal cell. They are used to cleave a separate nucleic acid, preferably
CC RNA. The nucleic acids are used to inhibit gene expression and/or cell
CC proliferation, and can be used to treat a disease or condition. More
CC than one nucleic acid can be independently targeted to the same or
CC different sites in a cell. The nucleic acids may be used to study DNA.
CC The modifications to the nucleic acids optimises their catalytic activity
CC and can maintain or enhance their activity. They exhibit a high degree
CC of specificity for RNA. The present sequence represents the coding
CC sequence of deoxyribozyme #8 used in the method of the invention.
XX
SQ Sequence 18 BP; 4 A; 4 C; 6 G; 4 U; 0 other;

Query Match 100.0%; Score 18; DB 22; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aauggcuaucgugcgca 18
Db 1 aauggcuaucgugcgca 18

RESULT 2
AAS12381
ID AAS12381 standard; RNA; 29 BP.
XX
AC AAS12381;
XX
DT 21-NOV-2001 (first entry)
XX
DE Class IV ribozyme.
XX
KW Deoxyribozyme; cytosstatic; endonuclease; RNA cleavage; DNA cleavage;
KW gene therapy; plant; fungus; bacteria; mammal; ribozyme; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_binding 1..8
FT /tag= a
FT /note= "Forms double-stranded region with bases 15
FT to 8 of AAS12374"
FT misc_binding 25..29
FT /tag= b
FT /note= "Forms double-stranded region with bases 5
FT to 1 of AAS12374"
XX
PN WO200159102-A2.
XX
PD 16-AUG-2001.
XX
PF 08-FEB-2001; 2001WO-US04223.
XX
PR 08-FEB-2000; 2000US-0181360.
PR 31-MAR-2000; 2000US-0193646.
XX
PA (RIBO-) RIBOZYME PHARM INC.
PA (UYVA) UNIV YALE.
XX
XX Breaker R, Beigelman L, Emilsson G;
XX WPI; 2001-536526/59.
XX
PT New nucleic acids with endonuclease activity, such as ribozymes and

PT nucleozymes, for modulating gene expression in a plant, mammalian,
PT bacterial or fungal cell -
PS Example 1; Fig 9; 96pp; English.
XX
XX
CC The invention relates to nucleic acid molecules with endonuclease
CC activity, which are particularly useful for cleavage of RNA or DNA.
CC The nucleic acids are used in a pharmaceutical composition and are used
CC to modulate expression of a gene in a plant, mammalian, bacterial or
CC fungal cell. They are used to cleave a separate nucleic acid, preferably
CC RNA. The nucleic acids are used to inhibit gene expression and/or cell
CC proliferation, and can be used to treat a disease or condition. More
CC than one nucleic acid can be independently targeted to the same or
CC different sites in a cell. The nucleic acids may be used to study DNA.
CC The modifications to the nucleic acids optimises their catalytic activity
CC and can maintain or enhance their activity. They exhibit a high degree
CC of specificity for RNA. The present sequence represents the Class IV
CC ribozyme, used in an example which demonstrates the method of
CC the invention.
XX
SQ Sequence 29 BP; 6 A; 7 C; 11 G; 5 U; 0 other;

Query Match 100.0%; Score 18; DB 22; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aauggcuaucgugcgca 18
Db 8 aauggcuaucgugcgca 25

RESULT 3
AAV99664/C
ID AAV99664 standard; DNA; 249 BP.
XX
AC AAV99664;
XX
DT 29-MAR-1999 (first entry)
XX
DE Vector M13trueblue modified lacZ alpha gene region.
XX
KW Beta-galactosidase; alpha-peptide; selectable marker;
KW marker inactivation; lacZ-alpha; vector; M13trueblue; ds; ss;
KW cyclic; circular.
XX
OS Chimeric - Escherichia coli.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Promoter 1..29
FT /tag= a
FT /note= "lac promoter"
FT RBS 35..45
FT /tag= b
FT CDS 49..273
FT /tag= d
FT /product= alpha peptide aa 1-54
FT 73..153124..198
FT /tag= e
FT /note= "Colour selection cloning sites"
XX
PN WO9850566-A1.
XX
PD 12-NOV-1998.
XX
PF 01-MAY-1998; 98WO-US08854.
XX
PR 07-MAY-1997; 97US-0852834.
XX
PA (LEBE/) LEBEL S.
PA (SLIL/) SLILATY S N.
XX

```
PI Label S, Sllaty SN;
XX
DR WPI: 1999-070096/06.
DR P-PSDB: AAW87835.
XX
PT New cloning vector containing lacZ alpha-peptide sequence with
PT cloning sites at specific positions - such that inserting DNA at
PT these sites generates few false negatives in subsequent colour
PT selection, e.g. for gap-free shotgun cloning and genomic library
PT development
XX
PS Example 3; Fig 3B; 62pp; English.
XX
CC This is the nucleotide sequence of the region of new cloning vector
CC M31trueblue (6974 bp) containing a modified Escherichia coli lacZ
CC alpha peptide gene sequence. Beta-galactosidase (bgal) lacZ alpha
CC (see AAW87788) gene fragments have been modified for use in new
CC cloning systems that use marker inactivation for the identification
CC of recombinants. The systems are based on the observation that
CC reliable inactivation of lacZ alpha occurs only if DNA is inserted
CC in the gene region encoding amino acids 8-38 of bgal. Claimed
CC cloning vectors comprise at least one promoter linked to a modified
CC lacZ alpha coding sequence containing at least one restriction site
CC introduced downstream of, and including, the codon for amino acid 8
CC of bgal. The high accuracy of colour selection afforded by the
CC modified lacZ alpha coding sequence allows the vector to be used
CC for general cloning purposes, for gap-free shotgun sequencing, for
CC facilitating industrial applications of gene isolation and genetic
CC engineering, and for development of ordered genomic libraries.
XX
SQ Sequence 249 BP; 60 A; 71 C; 67 G; 51 T; 0 other;

Query Match          77.8%; Score 14; DB 20; Length 249;
Best Local Similarity 78.6%; Pred. No. 8.4;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 ggcuaucgugcg 17
   ||||:||||:||||
DB 188 GGCTATCGGTGCG 175

RESULT 4
AAC61661/c
ID AAC61661 standard; DNA; 249 BP.
XX
AC AAC61661;
XX
DT 19-FEB-2001 (first entry)
XX
DE Nucleotide sequence of a modified lacZ N-terminal sequence.
XX
KM Beta-galactosidase; lacZ; alpha peptide; promoter; cloning system; ss.
XX
OS Synthetic.
OS Escherichia coli.
PN US6127171-A.
XX
PD 03-OCT-2000.
XX
PF 01-MAY-1998; 980US-0070842.
XX
PR 07-MAY-1997; 97US-0852834.
XX
PA (GENO-) GENOMICS ONE CORP.
XX
PI Label S, Sllaty SN;
XX
DR WPI: 2000-611058/58.
XX
PT Cloning systems with marker inactivation for identification of
PT recombinants with insertion of a polynucleotide, comprises a promoter
```

```
PT linked to modified lacZ alpha-gene and a restriction enzyme cleavable
PT cloning site
XX
PS Example 3; Fig 3B; 23pp; English.
XX
CC The present sequence represents a fragment of a modified lacZ sequence.
CC The fragment encodes the N-terminal of beta-galactosidase. The present
CC sequence was modified to contain multiple restriction enzyme sites.
CC A defective beta-galactosidase can be complemented by an alpha peptide.
CC The specification describes a vector comprising a promoter operably
CC linked to a modified Escherichia coli lacZ coding sequence encoding an
CC alpha peptide and containing at least a cloning site cleavable by a
CC restriction enzyme. The vector is useful for cloning a DNA molecule.
CC The vector is useful as a cloning system for the identification of
CC recombinants containing the insertion of a nucleic acid molecule.
XX
SQ Sequence 249 BP; 60 A; 71 C; 67 G; 51 T; 0 other;

Query Match          77.8%; Score 14; DB 21; Length 249;
Best Local Similarity 78.6%; Pred. No. 8.4;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 ggcuaucgugcg 17
   ||||:||||:||||
DB 188 GGCTATCGGTGCG 175

RESULT 5
AAV99663/c
ID AAV99663 standard; DNA; 372 BP.
XX
AC AAV99663;
XX
DT 29-MAR-1999 (first entry)
XX
DE Vector pTrueBlue modified lacZ alpha gene region.
XX
KM Beta-galactosidase; alpha-peptide; selectable marker;
KM marker inactivation; lacZ-alpha; vector; pTrueBlue; ds; cyclic;
KM circular.
XX
OS Chimeric - Escherichia coli.
OS Chimeric - bacteriophage T7.
OS Synthetic.
XX
FH Key
FH Location/Qualifiers
FT 1..29
FT /tag= a
FT /note= "lac promoter"
FT 35..45
FT /tag= b
FT /note= "Optimised ribosome binding site"
FT 58..75
FT /tag= c
FT /note= "T7 promoter"
FT 109..273
FT /tag= d
FT 124..198
FT /tag= e
FT /note= "colour selection cloning sites"
FT 287..297
FT /tag= f
FT /note= "ExoIII protection sites"
FT 317..346
FT /tag= g
FT /tag= 9
FT complement (349..372)
FT /tag= h
FT /note= "Ori fl"
XX
PN WO9850566-A1.
XX
PD 12-NOV-1998.
```


PF	01-MAY-1988:	98WO-US08854.
XX		
PR	07-MAY-1997:	97US-0852834.
XX		
PA	(LEBE/) LEBEL S.	
XX	(SLIL/) SLILATY S N.	
PI	Lebel S, Sillyaty SN;	
XX		
DR	WPI: 1999-070096/06.	
XX	P-PSDB; AAM87854.	
PT	New cloning vector containing lacZ alpha-peptide sequence with	
PR	cloning sites at specific positions - such that inserting DNA at	
PT	these sites generates few false negatives in subsequent colour	
PT	selection, e.g. for gap-free shotgun cloning and genomic library	
PT	development	
XX		
PS	Example 2; Fig 2B; 62pp; English.	
XX		
CC	This is the nucleotide sequence of the region of new cloning vector	
CC	pTrieblue (2914 bp) containing a modified Escherichia coli lacZ	
CC	alpha peptide gene sequence. Beta-galactosidase (bgal) lacZ alpha	
CC	(see AAM87788) gene fragments have been modified for use in new	
CC	cloning systems that use marker inactivation for the identification	
CC	of recombinants. The systems are based on the observation that	
CC	reliable inactivation of lacZ alpha occurs only if DNA is inserted	
CC	in the gene region encoding amino acids 8-38 of bgal. Claimed	
CC	cloning vectors comprise at least one promoter linked to a modified	
CC	lacZ alpha coding sequence containing at least one restriction site	
CC	introduced downstream of, and including, the codon for amino acid 8	
CC	of bgal. The high accuracy of colour selection afforded by the	
CC	modified lacZ alpha coding sequence allows the vector to be used	
CC	for general cloning purposes, for gap-free shotgun sequencing, for	
CC	facilitating industrial applications of gene isolation and genetic	
CC	engineering, and for development of ordered genomic libraries.	
SQ	Sequence 372 BP; 92 A; 91 C; 95 G; 94 T; 0 other;	
QY	4 ggcuaacgagucgc 17	
	: : :	
Db	233 GGCGTAFCGTGCG 220	
RESULT	6	77.8%; Score 14; DB 20; Length 372;
AAC61660/c	ID AAC61660 standard; DNA; 372 BP.	
XX	AAC61660:	
AC	19-FEB-2001 (first entry)	
XX		
DT	Nucleotide sequence of a modified lacZ N-terminal sequence.	
XX		
DE	Beta-galactosidase; lacZ; alpha peptide; promoter; cloning system; ss.	
XX		
OS	Synthetic.	
XX	Escherichia coli.	
PN	US6127171-A.	
PD	03-OCT-2000.	
XX		
PE	01-MAY-1998: 98US-0070842.	
XX		
PR	07-MAY-1997: 97US-0852834.	
XX		

PA	(GENO-)	GENOMICS ONE CORP.
XX		
PI	Label S,	Siliaty SN;
XX		
DR	WPI:	2000-611058/58.
XX		
PT	Cloning systems with marker inactivation for identification of	
PT	recombinants with insertion of a polynucleotide, comprises a promoter	
PT	linked to modified lacZ alpha-gene and a restriction enzyme cleavable	
PT	cloning site	
XX		
PS	Example 2; Fig 2B; 23pp;	English.
XX		
CC	The present sequence represents a fragment of a modified lacZ sequence.	
CC	The fragment encodes the N-terminal of beta-galactosidase. The present	
CC	sequence was modified to contain multiple restriction enzyme sites.	
CC	A defective beta-galactosidase can be complemented by an alpha peptide.	
CC	The specification describes a vector comprising a promoter operably	
CC	linked to a modified Escherichia coli lacZ coding sequence encoding an	
CC	alpha peptide and containing at least a cloning site cleavable by a	
CC	restriction enzyme. The vector is useful for cloning a DNA molecule.	
CC	The vector is useful as a cloning system for the identification of	
CC	recombinants containing the insertion of a nucleic acid molecule.	
XX		
SO	Sequence 372 BP; 92 A; 91 C; 95 G; 94 T; 0 other;	
	Query Match	77.8%; Score 14; DB 21; Length 372;
	Best Local Similarity	78.6%; Pred. No. 8.4;
	Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;	
OY	4 ggcuaucgugucg 17	
	: :	
Db	233 GGCTATCGTCCG 220	
	RESULT 7	
ID	AAV99665/c	
XX	AAV99665 standard; DNA: 375 BP.	
XX		
AC	AAV99665;	
XX		
DT	29-MAR-1999 (first entry)	
XX		
DE	Vector pTrueBlue-Bac modified lacZ alpha gene region.	
XX		
XX	Beta-galactosidase; alpha-peptide; selectable marker;	
KM	marker inactivation; lacZ-alpha; vector; pTrueBlue-Bac; ds; ss;	
KM	cyclic; circular.	
XX		
OS	Chimeric - Escherichia coli.	
OS	Chimeric - bacteriophage T7.	
OS	Synthetic.	
XX		
PH		
FT	Key	Location/Qualifiers
FT	promoter	1..29
FT		/tag= a
FT		/note= "lac promoter"
FT	RBS	35..45
FT		/tag= b
FT		/note= "optimised ribosome binding site"
FT		58..75
FT	promoter	
FT		/tag= c
FT		/note= "T7 promoter"
FT		109..273
FT	CDS	
FT		/tag= d
FT		124..198
FT	misc_feature	
FT		/tag= e
FT		/note= "colour selection cloning sites"
FT	misc_feature	287..297
FT		/tag= f
FT		/note= "ExoIII protection sites"
FT		317..346
FT	terminator	

```
FT misc_feature /*tag= 9
FT complement (349..372)
FT /*tag= h
FT /*note= "Ori fl"
XX
XX
XX WO9850566-A1.
XX
XX 12-NOV-1998.
XX
XX 01-MAY-1998; 98WO-US08854.
XX
XX 07-MAY-1997; 97US-0852834.
XX
XX (LEBE/) LEBEL S.
XX (SLI/) SLIATY S N.
XX
XX Lebel S, Sliaty SN;
XX
XX WPI; 1999-070096/06.
XX P-PSDB; AAW87854.
XX
XX
XX New cloning vector containing lacZ alpha-peptide sequence with
XX cloning sites at specific positions - such that inserting DNA at
XX these sites generates few false negatives in subsequent colour
XX selection, e.g. for gap-free shotgun cloning and genomic library
XX development
XX
XX Example 4; Fig 4B; 62pp; English.
XX
XX This is the nucleotide sequence of the region of new cloning vector
XX pTrieBlue-bac (7289 bp) containing a modified Escherichia coli lacZ
XX alpha peptide gene sequence. Beta-galactosidase (bgal) lacZ alpha
XX (see AAW87788) gene fragments have been modified for use in new
XX cloning systems that use marker inactivation for the identification
XX of recombinants. The systems are based on the observation that
XX reliable inactivation of lacZ alpha occurs only if DNA is inserted
XX in the gene region encoding amino acids 8-38 of bgal. Claimed
XX cloning vectors comprise at least one promoter linked to a modified
XX lacZ alpha coding sequence containing at least one restriction site
XX introduced downstream of, and including, the codon for amino acid 8
XX of bgal. The high accuracy of colour selection afforded by the
XX modified lacZ alpha coding sequence allows the vector to be used
XX for general cloning purposes, for gap-free shotgun sequencing, for
XX facilitating industrial applications of gene isolation and genetic
XX engineering, and for development of ordered genomic libraries.
XX
XX Sequence 375 BP; 92 A; 96 C; 98 G; 89 T; 0 other.
XX
XX
XX Query Match 77.8%; Score 14; DB 20; Length 375;
XX Best Local Similarity 78.6%; Pred. No. 8.4;
XX Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 4 ggcctatcggtcgcg 17
XX ||||:||||:||||
XX Db 233 GGCCTATCGGTGCG 220
XX
XX RESULT 8
XX AAC61662/C
XX ID AAC61662 standard; DNA; 375 BP.
XX
XX AC AAC61662;
XX
XX 19-FEB-2001 (first entry)
XX
XX Nucleotide sequence of a modified lacZ N-terminal sequence.
XX
XX DE Beta-galactosidase; lacZ; alpha peptide; promoter; cloning system; ss.
XX
XX KM Synthetic.
XX
XX OS Escherichia coli.
XX
XX
```

```
PN US6127171-A.
XX
XX 03-OCT-2000.
XX
XX 01-MAY-1998; 98US-0070842.
XX
XX 07-MAY-1997; 97US-0852834.
XX
XX (GENO-) GENOMICS ONE CORP.
XX
XX Lebel S, Sliaty SN;
XX
XX WPI; 2000-611058/58.
XX
XX
XX Cloning systems with marker inactivation for identification of
XX recombinants with insertion of a polynucleotide, comprises a promoter
XX linked to modified lacZ alpha-gene and a restriction enzyme cleavable
XX cloning site
XX
XX Example 4; Fig 4B; 23pp; English.
XX
XX The present sequence represents a fragment of a modified lacZ sequence.
XX The fragment encodes the N-terminal of beta-galactosidase. The present
XX sequence was modified to contain multiple restriction enzyme sites.
XX A defective beta-galactosidase can be complemented by an alpha peptide.
XX The specification describes a vector comprising a promoter operably
XX linked to a modified Escherichia coli lacZ coding sequence encoding an
XX alpha peptide and containing at least a cloning site cleavable by a
XX restriction enzyme. The vector is useful for cloning a DNA molecule.
XX The vector is useful as a cloning system for the identification of
XX recombinants containing the insertion of a nucleic acid molecule.
XX
XX Sequence 375 BP; 92 A; 96 C; 98 G; 89 T; 0 other.
XX
XX
XX Query Match 77.8%; Score 14; DB 21; Length 375;
XX Best Local Similarity 78.6%; Pred. No. 8.4;
XX Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 4 ggcctatcggtcgcg 17
XX ||||:||||:||||
XX Db 233 GGCCTATCGGTGCG 220
XX
XX RESULT 9
XX AAAS3641
XX ID AAAS3641 standard; DNA; 3839 BP.
XX
XX AC AAAS3641;
XX
XX 04-DEC-2000 (first entry)
XX
XX TT virus (TTV-US35) genomic sequence.
XX
XX TTV: TT virus; blood transmission; detection; amplification; primer;
XX transplantation; xenotransplantation; vector; ss.
XX
XX TT virus isolate US35.
XX
XX WO200046407-A2.
XX
XX 10-AUG-2000.
XX
XX 04-FEB-2000; 2000WO-US02982.
XX
XX 05-FEB-1999; 99US-0245248.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Leary TP, Simons JN, Erker JC, Chalmers ML, Birkenmeyer LG;
XX Puerhoff AS, Pilot-Matias TJ, Desai SM, Mushahwar IK;
XX
XX WPI; 2000-514969/46.
XX
```

XX New oligomer primer useful for the detection of TT virus in test
PT samples and tissues and organs for use in (xeno)transplantation
XX
PS Claim 1: Page 113-115; 139pp; English.
XX
CC Hundreds of TT virus (TTV) isolates have been identified by PCR assays,
CC which have amplified less than 400 nucleotides of sequence. Only one
CC full length TTV (Isolate GH1 - see AAK53632) and two near full length
CC isolates (TA278 and TTV CHN1) have been reported. Therefore, in an
CC attempt to more fully understand the TTV genome, several divergent
CC isolates of TTV were extended to full or near full length (shown in
CC AAK53637-44). These sequences revealed up to 30 percent nucleotide
CC divergence, 3 conserved open reading frames, a lack of identifiable
CC regulatory elements, and the presence of distinct genotypes and subtypes.
CC TTV is a circular, negative single-stranded DNA virus. Isolate GH1 was
CC 3852 nucleotides in length, 113 nucleotides longer than previously
CC reported. The newly discovered region is GC rich (89 percent) and
CC contains several potential stem-loop structures. TTV DNA can be
CC transmitted by blood or blood products. It is also possible that TTV is
CC transmitted by a faecal-oral route, demonstrated by the presence of TTV
CC in the faeces of infected humans. Detection of TTV in test samples can be
CC enhanced by use of DNA amplification assays that use DNA oligomers as
CC primers. The primers are useful for detecting the presence of TTV target
CC nucleotides in biological samples and tissues and organs to be used in
CC transplantation and xenotransplantation (claimed). The TTV genome itself
CC can be used as a vector in order to introduce heterologous DNA into a
CC host cell.
XX
SQ Sequence 3839 BP; 1082 A; 1075 C; 920 G; 762 T; 0 other;
XX
Query Match 77.8%; Score 14; DB 21; Length 3839;
Best Local Similarity 71.4%; Pred. No. 7.9;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 2 auggcacuaacgagug 15
1:||||:|||||
Db 579 atggcctatcggtg 592
RESULT 10
AAK55950/C
ID AAK55950 standard; cDNA; 415 BP.
XX
AC AAK55950;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:1010.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytosolic; gene therapy; vaccine; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217486.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0233400.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234222.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0234984.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235835.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
DR P-PSDB; AAM83169.
XX
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
PS Claim 1; SEQ ID NO 1010; 3071bp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome

CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK54703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention.
XX
SO Sequence 415 BP; 131 A; 101 C; 82 G; 98 T; 3 other;

Query Match 72.2%; Score 13; DB 22; Length 415;
Best Local Similarity 76.9%; Pred. No. 36;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaagccuacug 13
11:||||:||||
DB 272 AATGGCCTATCGG 260

RESULT 11
AAFO9281
ID AAF09281 standard; cDNA; 652 BP.
XX
AC AAF09281;
XX
DT 13-MAR-2001 (first entry)
XX
DE Fusarium venenatum EST SEQ ID NO:1804.
XX
XX
XX Multiple gene expression: filamentous fungal cell; EST.
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS Fusarium venenatum.
XX
PN WO200056762-A2.
XX
PD 28-SEP-2000.
XX
PF 22-MAR-2000; 2000WO-US07781...
XX
XX 22-MAR-1999; 99US-0273623.
XX
XX (NOVO) NOVO NORDISK BIOTECH INC.
PA (NOVO) NOVO NORDISK AS.
XX
XX Berka RM, Rey MW, Shuster JR, Kaupinen S, Clausen IG, Olsen PB;
PI WPI; 2000-594572/56.
XX
XX
XX Monitoring differential expression of genes in filamentous fungal cells
PT uses fluorescence-labeled nucleic acids isolated from the cells and a
PT substrate of expressed sequence tags -
XX
XX
PS Claim 86; Page 1058-1059; 3161bp; English.
XX
XX The present invention describes a method for monitoring differential
CC expression of genes in a first filamentous fungal (FF) cell relative to
CC expression of the same genes in one or more second filamentous fungal
CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes
CC in a first filamentous fungal (FF) cell relative to expression of the
CC same genes in one or more second filamentous fungal cells. Monitoring
CC the global expression of genes from FF cells allows the production
CC potential of the microorganisms to be improved. New genes may be
CC discovered, possible functions of unknown open reading frames can be

CC identified and gene copy number variation and stability can be
CC monitored. The expression of genes can be used to study how FR cells
CC adapt to changes in culture conditions, environmental stress, spore
CC morphogenesis, recombination, metabolic or catabolic pathway
CC engineering. Using ESTs provides several advantages over genomic or
CC random cDNA clones including elimination of redundancy as one spot on an
CC array equals one gene or open reading frame, and organisation of the
CC microarrays based on function of the gene products to facilitate
CC analysis of the results. AAF07478 to AAF11247 represents ESTs from
CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
CC niger; AAF11854 to AAF14578 represents ESTs from Aspergillus oryzae; and
CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
CC all specifically claimed in the present invention.

SO Sequence 652 BP; 155 A; 224 C; 124 G; 144 T; 5 other;

Query Match 72.2%; Score 13; DB 21; Length 652;
Best Local Similarity 69.2%; Pred. No. 36;
Matches 9; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 auggccuacggu 14
1:||||:||||:
Db 575 atggcctatcggt 587

RESULT 12
AAH67911/c
ID AAH67911 standard; DNA; 1011 BP.

XX AAH67911;

DT 26-SEP-2001 (first entry)

XX C glutamicum coding sequence fragment SEQ ID NO: 2946.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

KW organic acid synthesis; ds.

XX Corynebacterium glutamicum.

OS EPI108790-A2.

XX EPI108790-A2.

PD 20-JUN-2001.

XX 20-JUN-2001.

PF 18-DEC-2000; 2000EP-0127688.

XX 18-DEC-2000; 2000EP-0127688.

PR 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.

PR 03-AUG-2000; 2000JP-0280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR WPI; 2001-376931/40.

XX P-PSDB; AAG92692.

PT Novel polynucleotides derived from Coryneform bacteria, for identifying

PT mutation point of a gene, measuring expression of a gene, analysing

PT expression profile or pattern of a gene and identifying homologous gene

PS Claim 8: SEQ ID NO: 2946; 246pp + Sequence Listing: English.

CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.

XX

SO Sequence 1011 BP; 250 A; 280 C; 260 G; 221 T; 0 other;

Query Match 72.2%; Score 13; DB 22; Length 1011;
Best Local Similarity 76.9%; Pred. No. 36;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 auggccuacggu 13
1:||||:||||:
Db 371 AATGACCTATCGG 359

RESULT 13
AAF71301/c
ID AAF71301 standard; DNA; 1053 BP.

XX AAF71301;

DT 30-APR-2001 (first entry)

XX Corynebacterium glutamicum HA protein nucleotide sequence SEQ ID NO:327.

KW Corynebacterium glutamicum; homeostasis; adaptation; HA protein;

KW fine chemical production; organic acid; proteanogenic amino acid;

KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;

KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;

KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;

KW diagnosis; Corynebacterium diptheriae; genetic engineering;

XX Brevibacterium; environmental condition; ds.

XX Corynebacterium glutamicum.

OS WO200100842-A2.

XX WO200100842-A2.

PD 04-JAN-2001.

XX 04-JAN-2001.

PF 23-JUN-2000; 2000WO-IB00911.

XX 23-JUN-2000; 2000WO-IB00911.

PR 25-JUN-1999; 99US-0141031.

PR 08-JUL-1999; 99DE-1031636.

PR 09-JUL-1999; 99DE-1032125.

PR 09-JUL-1999; 99DE-1032126.

PR 09-JUL-1999; 99DE-1032127.

PR 09-JUL-1999; 99DE-1032128.

PR 09-JUL-1999; 99DE-1032129.

PR 09-JUL-1999; 99DE-1032226.

PR 14-JUL-1999; 99DE-1032920.

PR 14-JUL-1999; 99DE-1032922.

XX 14-JUL-1999; 99DE-1032924.

CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.

XX

SO Sequence 1011 BP; 250 A; 280 C; 260 G; 221 T; 0 other;

Query Match 72.2%; Score 13; DB 22; Length 1011;
Best Local Similarity 76.9%; Pred. No. 36;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 auggccuacggu 13
1:||||:||||:
Db 371 AATGACCTATCGG 359

RESULT 13
AAF71301/c
ID AAF71301 standard; DNA; 1053 BP.

XX AAF71301;

DT 30-APR-2001 (first entry)

XX Corynebacterium glutamicum HA protein nucleotide sequence SEQ ID NO:327.

KW Corynebacterium glutamicum; homeostasis; adaptation; HA protein;

KW fine chemical production; organic acid; proteanogenic amino acid;

KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;

KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;

KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;

KW diagnosis; Corynebacterium diptheriae; genetic engineering;

XX Brevibacterium; environmental condition; ds.

XX Corynebacterium glutamicum.

OS WO200100842-A2.

XX WO200100842-A2.

PD 04-JAN-2001.

XX 04-JAN-2001.

PF 23-JUN-2000; 2000WO-IB00911.

XX 23-JUN-2000; 2000WO-IB00911.

PR 25-JUN-1999; 99US-0141031.

PR 08-JUL-1999; 99DE-1031636.

PR 09-JUL-1999; 99DE-1032125.

PR 09-JUL-1999; 99DE-1032126.

PR 09-JUL-1999; 99DE-1032127.

PR 09-JUL-1999; 99DE-1032128.

PR 09-JUL-1999; 99DE-1032129.

PR 09-JUL-1999; 99DE-1032226.

PR 14-JUL-1999; 99DE-1032920.

PR 14-JUL-1999; 99DE-1032922.

XX 14-JUL-1999; 99DE-1032924.

CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.

XX

SO Sequence 1011 BP; 250 A; 280 C; 260 G; 221 T; 0 other;

Query Match 72.2%; Score 13; DB 22; Length 1011;
Best Local Similarity 76.9%; Pred. No. 36;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 auggccuacggu 13
1:||||:||||:
Db 371 AATGACCTATCGG 359

RESULT 13
AAF71301/c
ID AAF71301 standard; DNA; 1053 BP.

XX AAF71301;

DT 30-APR-2001 (first entry)

XX Corynebacterium glutamicum HA protein nucleotide sequence SEQ ID NO:327.

KW Corynebacterium glutamicum; homeostasis; adaptation; HA protein;

KW fine chemical production; organic acid; proteanogenic amino acid;

KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;

KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;

KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;

KW diagnosis; Corynebacterium diptheriae; genetic engineering;

XX Brevibacterium; environmental condition; ds.

XX Corynebacterium glutamicum.

OS WO200100842-A2.

XX WO200100842-A2.

PD 04-JAN-2001.

XX 04-JAN-2001.

PF 23-JUN-2000; 2000WO-IB00911.

XX 23-JUN-2000; 2000WO-IB00911.

PR 25-JUN-1999; 99US-0141031.

PR 08-JUL-1999; 99DE-1031636.

PR 09-JUL-1999; 99DE-1032125.

PR 09-JUL-1999; 99DE-1032126.

PR 09-JUL-1999; 99DE-1032127.

PR 09-JUL-1999; 99DE-1032128.

PR 09-JUL-1999; 99DE-1032129.

PR 09-JUL-1999; 99DE-1032226.

PR 14-JUL-1999; 99DE-1032920.

PR 14-JUL-1999; 99DE-1032922.

XX 14-JUL-1999; 99DE-1032924.

CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.

XX

SO Sequence 1011 BP; 250 A; 280 C; 260 G; 221 T; 0 other;

Query Match 72.2%; Score 13; DB 22; Length 1011;
Best Local Similarity 76.9%; Pred. No. 36;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 auggccuacggu 13
1:||||:||||:
Db 371 AATGACCTATCGG 359

RESULT 13
AAF71301/c
ID AAF71301 standard; DNA; 1053 BP.

XX AAF71301;

DT 30-APR-2001 (first entry)

XX Corynebacterium glutamicum HA protein nucleotide sequence SEQ ID NO:327.

KW Corynebacterium glutamicum; homeostasis; adaptation; HA protein;

KW fine chemical production; organic acid; proteanogenic amino acid;

KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;

KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;

KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;

KW diagnosis; Corynebacterium diptheriae; genetic engineering;

XX Brevibacterium; environmental condition; ds.

XX Corynebacterium glutamicum.

OS WO200100842-A2.

XX WO200100842-A2.

PD 04-JAN-2001.

XX 04-JAN-2001.

PF 23-JUN-2000; 2000WO-IB00911.

XX 23-JUN-2000; 2000WO-IB00911.

PR 25-JUN-1999; 99US-0141031.

PR 08-JUL-1999; 99DE-1031636.

PR 09-JUL-1999; 99DE-1032125.

PR 09-JUL-1999; 99DE-1032126.

PR 09-JUL-1999; 99DE-1032127.

PR 09-JUL-1999; 99DE-1032128.

PR 09-JUL-1999; 99DE-1032129.

PR 09-JUL-1999; 99DE-1032226.

PR 14-JUL-1999; 99DE-1032920.

PR 14-JUL-1999; 99DE-1032922.

XX 14-JUL-1999; 99DE-1032924.

CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.

XX

SO Sequence 1011 BP; 250 A; 280 C; 260 G; 221 T; 0 other;

Query Match 72.2%; Score 13; DB 22; Length 1011;
Best Local Similarity 76.9%; Pred. No. 36;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 auggccuacggu 13
1:||||:||||:
Db 371 AATGACCTATCGG 359

RESULT 13
AAF71301/c
ID AAF71301 standard; DNA; 1053 BP.

XX AAF71301;

DT 30-APR-2001 (first entry)

XX Corynebacterium glutamicum HA protein nucleotide sequence SEQ ID NO:327.

KW Corynebacterium glutamicum; homeostasis; adaptation; HA protein;

KW fine chemical production; organic acid; proteanogenic amino acid;

KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;

KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;

KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;

KW diagnosis; Corynebacterium diptheriae; genetic engineering;

XX Brevibacterium; environmental condition; ds.

XX Corynebacterium glutamicum.

OS WO200100842-A2.

XX WO200100842-A2.

PD 04-JAN-2001.

XX 04-JAN-2001.

PF 23-JUN-2000; 2000WO-IB00911.

XX 23-JUN-2000; 2000WO-IB00911.

PR 25-JUN-1999; 99US-0141031.

PR 08-JUL-1999; 99DE-1031636.

PR 09-JUL-1999; 99DE-1032125.

PR 09-JUL-1999; 99DE-1032126.

PR 09-JUL-1999; 99DE-1032127.

PR 09-JUL-1999; 99DE-1032128.

PR 09-JUL-1999; 99DE-1032129.

PR 09-JUL-1999; 99DE-1032226.

PR 14-JUL-1999; 99DE-1032920.

PR 14-JUL-1999; 99DE-1032922.

XX 14-JUL-1999; 99DE-1032924.

CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.

XX

SO Sequence 1011 BP; 250 A; 280 C; 260 G; 221 T; 0 other;

Query Match 72.2%; Score 13; DB 22; Length 1011;
Best Local Similarity 76.9%; Pred. No. 36;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 auggccuacggu 13
1:||||:||||:
Db 371 AATGACCTATCGG 359

RESULT 13
AAF71301/c
ID AAF71301 standard; DNA; 1053 BP.

XX AAF71301;

DT 30-APR-2001 (first entry)

XX Corynebacterium glutamicum HA protein nucleotide sequence SEQ ID NO:327.

KW Corynebacterium glutamicum; homeostasis; adaptation; HA protein;

KW fine chemical production; organic acid; proteanogenic amino acid;

KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;

KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;

KW carbohydrate; aromatic

PI Pompeius M, Kroeger B, Schroeder H, Zeider O, Haberhauer G;
XX WPI: 2001-061974/07.
DR P-PSDB; AAB79186.
XX
XX
PT New isolated Corynebacterium glutamicum nucleic acid for production or
PT modulation of production of fine chemicals such as amino acids,
PT nucleosides, nucleotides, lipids, fatty acids, carbohydrates, vitamins
PT or enzymes -
XX
PS Claim 3; Page 570-571; 712pp; English.
XX
CC AAF71138 to AAF71357 encode the Corynebacterium glutamicum homeostasis
CC and adaptation (HA) proteins given in AAB79023 to AAB79242. The
CC C. glutamicum HA genes (I) can be used in vectors for expression in host
CC cells and production of fine chemicals, such as, an organic acid,
CC proteinogenic or nonproteinogenic amino acid (preferred), putine or
CC pyrimidine base, nucleoside, nucleotide, lipid, saturated or unsaturated
CC fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,
CC polyketide or enzyme. The amino acids produced can be lysine, glutamine,
CC glutamate, alanine, aspartate, glycine, serine, threonine, methionine,
CC cysteine, valine, leucine, isoleucine, arginine, proline, histidine,
CC tyrosine, phenylalanine, or tryptophan. The fine chemical production can
CC be modulated. The presence of (I) or HA proteins encoded by them are
CC used for diagnosing the presence or activity of Corynebacterium
CC diptheriae. (I) can be used to map the C. glutamicum genome or can be
CC used as markers for genetically engineered Corynebacterium or
CC Brevibacterium. The HA proteins encoded by the (I) are used to maintain
CC homeostasis in C. glutamicum or help the microorganism to adapt to
CC different environmental conditions.
XX
SQ Sequence 1053 BP; 256 A; 289 C; 272 G; 236 T; 0 other;

Query Match 72.2%; Score 13; DB 22; Length 1053;
Best Local Similarity 76.9%; Pred. No. 36;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 aauggccuacg 13
||:||||:||||
DB 390 AATGCCCTATCGG 378

RESULT 14
ID ABA21153 standard; DNA; 1128 BP.
XX
XX ABA21153;
AC
DT 23-JAN-2002 (first entry)
XX
XX
DE Human nervous system related polynucleotide SEQ ID NO 13484.
XX
XX
KM Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KM immunosuppressive; antitumor; anti-HIV; antibacterial; vulnery;
KM antiparkinsonian; antischizoid; antianemic; antidiabetic; cancer;
KM antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KM antiallergic; antidiabetic; antidiabetic; anticonvulsant; antifungal;
KM antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KM neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200159063-A2.
PN
XX
XX 16-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US01334.
PF
XX
XX 31-JAN-2000; 2000US-0179065.
PR
XX 04-FEB-2000; 2000US-0180628.
PR
XX 24-FEB-2000; 2000US-0184664.
PR
XX 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225216.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226581.
PR 22-AUG-2000; 2000US-0226582.
PR 22-AUG-2000; 2000US-0226586.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231142.
PR 08-SEP-2000; 2000US-0231143.
PR 08-SEP-2000; 2000US-0231144.
PR 08-SEP-2000; 2000US-0231145.
PR 08-SEP-2000; 2000US-0231146.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234598.
PR 25-SEP-2000; 2000US-0234599.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249247.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX PT useful for preventing, diagnosing and/or treating nervous system
XX PT cancers and metastases -
XX
XX
XX Disclosure; SEQ ID NO 13484; 1701bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABAI1004-ABA21534) and proteins

CC (ABAI678-ABAI8001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 1128 BP; 294 A; 284 C; 293 G; 257 T; 0 other;

Query Match 72.28; Score 13; DB 22; Length 1128;
Best Local Similarity 76.94; Pred. No. 36;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 6 ccaucgagucga 18
||:|||||
Db 802 cctatcggtcga 814

RESULT 15
AAS30162/c
ID AAS30162 standard; DNA: 1128 BP.
XX
AC AAS30162;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human lung antigen genomic DNA #232.
XX
KW Lung antigen protein; human; mouse; rabbit; goat; horse; cat; dog;
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW antineumatic; antiproliferative; cytosolic; cardiac; neuroprotective;
KW cerebroprotective; nocrotropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnerray; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; food preservative;
KW tissue regeneration; anti-infertility; food additive.
XX
OS Homo sapiens.
XX
PN WO200155303-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01301.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234598.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235636.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236602.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0239835.
PR 13-OCT-2000; 2000US-0239837.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246539.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251889.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-457723/49.

Isolated polypeptide for treating, preventing and/or prognosing
respiratory disorders related to the lung including lung cancers and
also for testing and detection e.g. diagnosis -

Claim 1; SEQ ID NO 426; 507pp; English.

XX Sequences AAS29931-AAS30164 represent genomic DNA molecules, which encode
XX the lung antigen polypeptides of the invention. Lung antigen polypeptides
XX and their associated polynucleotides are useful in the diagnosis,
XX treatment and prevention of various types of disorders in e.g. humans,
XX mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A
XX pathological condition can be determined by detecting the presence or
XX absence of a mutation in a lung antigen polynucleotide. The treatable
XX disorders include autoimmune diseases such as rheumatoid arthritis,
XX hyperproliferative disorders such as neoplasms of the breast or liver,
XX cardiovascular disorders such as cardiac arrest, cerebrovascular

CC disorders such as cerebral ischaemia, nervous system disorders such as
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,
 CC ocular disorders such as corneal infection, endocrine disorders such as
 CC premature labour and infertility, gastrointestinal disorders such as
 CC Crohn's disease, renal disorders such as glomerulonephritis and
 CC respiratory disorders such as asthma and pleurisy. The polypeptides can
 CC also be used to aid wound healing, to prevent skin aging due to sunburn,
 CC to maintain organs before transplantation, to regenerate tissues and in
 CC chemotaxis. The polypeptides can also be used as a food additive or
 CC preservative to increase or decrease storage capabilities.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

Query Match 72.2%; Score 13; DB 22; Length 1128;
 Best Local Similarity 76.9%; Pred. No. 36;
 Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ccuauugugugca 18
 ||:|||||
 Db 327 CCTATCGGTGCA 315

Search completed: September 9, 2002, 01:51:29
 Job time: 4529 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 9, 2002, 01:46:37 ; Search time 1931.22 seconds
(without alignments)
195.046 Million cell updates/sec

Title: US-09-780-929-98

Perfect score: 18
Sequence: 1 aaagcccaucgugcgga 18

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size: 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_ph:*
7: gb_pl:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	18	100.0	18	AX214296	Sequence
2	15	100.0	29	AX214316	Sequence
3	18	83.3	35221	CER04C12	Sequence
4	14	77.8	92	AR111665	Sequence
5	14	77.8	249	AR111672	Sequence
6	14	77.8	372	AR111659	Sequence
7	14	77.8	375	AR111673	Sequence
8	14	77.8	1100	AC048161	Sequence
9	14	77.8	1746	X1707162	Sequence
10	14	77.8	2082	AF267298	Sequence
11	14	77.8	2437	BSSPC	Sequence
12	14	77.8	3839	AF122920	Sequence
13	14	77.8	5798	RER1XGHIS	Sequence
14	14	77.8	10759	BACRPLP	Sequence
15	14	77.8	98950	AC034257	Sequence
16	14	77.8	155909	AC097562	Sequence
17	14	77.8	157601	AP004662	Sequence
18	14	77.8	157695	AL136442	Sequence
19	14	77.8	171099	AP003871	Sequence
20	14	77.8	182375	AC105920	Sequence
21	14	77.8	194081	AC024990	Sequence
22	14	77.8	213080	BSUB0001	Sequence
23	14	77.8	226130	AC108892	Sequence
24	14	77.8	228181	AC108073	Sequence
25	14	77.8	258050	AT141453	Sequence
26	13	72.2	446	HSB323242	Sequence
27	13	72.2	653	AF302741	Sequence
28	13	72.2	654	AY008466	Sequence
29	13	72.2	685	HSB336233	Sequence
30	13	72.2	704	HSB336457	Sequence
31	13	72.2	828	AF031656	Sequence
32	13	72.2	933	SCRNA9	Sequence
33	13	72.2	933	YSCRG552	Sequence
34	13	72.2	1011	AX123030	Sequence
35	13	72.2	1053	AX065983	Sequence
36	13	72.2	1190	ARI70982	Sequence
37	13	72.2	1225	SCE243217	Sequence
38	13	72.2	1260	AF306495	Sequence
39	13	72.2	1562	AK001979	Sequence
40	13	72.2	1928	BC015725	Sequence
41	13	72.2	1933	AY071002	Sequence
42	13	72.2	2448	AB031867	Sequence
43	13	72.2	2448	AB031868	Sequence
44	13	72.2	2512	SCRDN01	Sequence
45	13	72.2	2554	SCRDN02	Sequence

ALIGNMENTS

RESULT 1	AX214296	18 bp	linear	PAT 06-SEP-2001
LOCUS	AX214296	Sequence 109 from Patent WO0159102.		
DEFINITION	AX214296			
ACCESSION	AX214296			
VERSION	AX214296.1	GI:15524373		
KEYWORDS				
SOURCE				
ORGANISM				
synthetic construct.				
artificial sequence.				
REFERENCE				
AUTHORS	1 (bases 1 to 18)			
TITLE	Breaker, R. and Emmlison, G.			
JOURNAL	Nucleozymes with endonuclease activity			
FEATURES	Patent: WO 0159102-A 109 16-AUG-2001;			
source	PHARMACEUTICALS, INC. (US) ; Yale University (US)			
	Location/Qualifiers			
	1..18			
	/organism="synthetic construct"			
	/db_xref="taxon:32630"			
	/note="Nucleic Acid"			

BASE COUNT	ORIGIN
------------	--------

4 a	4 c	6 g	4 t
-----	-----	-----	-----

Query Match 100.0%; Score 18; DB 6; Length 18;
Best Local Similarity 77.8%; Pred. No. 0.86;
Matches 14; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 aaugcccaugcgugcgca 18
11:11111:1:111:1111
Db 1 AATGGCCTATCGGTGCGA 18

RESULT 2
AX214316 29 bp mRNA linear PAT 06-SEP-2001
LOCUS Sequence 129 from Patent WO0159102.
ACCESSION AX214316
VERSION AX214316.1 GI:15524393
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 29)
AUTHORS Breaker,R. and Emilsson,G.
TITLE Nucleozymes with endonuclease activity
JOURNAL Patent: WO 0159102-A 129 16-AUG-2001;
RIBOZYME PHARMACEUTICALS, INC. (US) ; Yale University (US)
FEATURES
Source 1..29
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Nucleic Acid"

BASE COUNT 6 a 7 c 11 g 5 t
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 29;
Best Local Similarity 77.8%; Pred. No. 0.83;
Matches 14; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 aaugcccaugcgugcgca 18
11:11111:1:111:1111
Db 8 AATGGCCTATCGGTGCGA 25

RESULT 3
CE704C12 35221 bp DNA linear INV 11-DEC-2001
LOCUS Caenorhabditis elegans cosmid T04C12, complete sequence.
ACCESSION Z81584
VERSION Z81584.1 GI:1914459
KEYWORDS HTG.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
REFERENCE 1 (sites)
AUTHORS none.
TITLE Genome sequence of the nematode C. elegans: a platform for
JOURNAL investigating biology. The C. elegans Sequencing Consortium
MEDLINE Science 282 (5396), 2012-2018 (1998)
REMARK 99069613
AUTHORS The C. elegans Sequencing Consortium.
TITLE 2 (bases 1 to 35221)
JOURNAL McMurtry,A.A.
DIRECT SUBMISSION
SUBMITTED (06-NOV-1996) Nematode Sequencing Project, Sanger
Institute, Hinxton, Cambridge CB10 1SA, England and Department of
Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
jes.sanger.ac.uk or rwenematode.wustl.edu
On Mar 30, 1997 this sequence version replaced gi:1695088.
Coding sequences below are predicted from computer analysis, using
predictions from GeneFinder (P. Green, U. Washington), and other
available information.
Current sequence finishing criteria for the C. elegans genome

sequencing consortium are that all bases are either sequenced
unambiguously on both strands, or on a single strand with both a
dye primer and dye terminator reaction, from distinct subclones.
Exceptions are indicated by an explicit note.
This sequence is the entire insert of clone T04C12. The true right
end of clone C45B11 is at 29588 in this sequence. The start of this
sequence (1..104) overlaps with the end of sequence Z74029.
The end of this sequence (35117..35221) overlaps with the start of
sequence AL032618.
For a graphical representation of this sequence and its analysis
see: [http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?](http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=T04C12)
IMPORTANT: This sequence is NOT necessarily the entire insert of
the specified clone. It may be shorter because we only sequence
overlapping sections once, or longer because we arrange for a small
overlap between neighbouring submissions.

FEATURES
Source 1..35221
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/chromosome="v"
/clone="T04C12"
complement(join(3712..3989,4068..4161,4243..4433,
4481..4626,4841..5058))
/gene="T04C12.2"
complement(join(3712..3989,4068..4161,4243..4433,
4481..4626,4841..5058))
/note="contains similarity to Pfam domain: PF01604 (7TM
chemoreceptor), Score=100.5, E-value=1.1e-26, N=1"
/codon_start=1
/protein_id="CA04680.1"
/db_xref="GI:3879479"
/db_xref="SPTREMBL:002296"
/translation="MTSTNTLYYSNKKKCKSDSFLASMOGLSVESHMLVEFPI
YGFYTCILDKPTKTNNSYKVLNTHCNCQYDILCSITPYEPPTSGPVGIL
RVLKVPSTVOLIGFISALFMAISVALFENRSADIONNFRITKKRMLIYSYKCF
IVLVYLIPPCNVPEODESARLHLQALPCPEFEFYSDVWIDKFWIYLVWMTAI
MISSNTRKFORSEFLGITQAVVLLFLPLFVLGIYVCEYVNOELNNSIVLFLSL
HGFTSFVITLVHHPYRRLIKVYTFERSAGK"
complement(join(6609..6736,7001..7177,7543..7583,
7626..7740,7998..8190))
/gene="T04C12.1"
complement(join(6609..6736,7001..7177,7543..7583,
7626..7740,7998..8190))
/note="cDNA EST yk656b12.3 comes from this gene"
cDNA EST yk656b12.5 comes from this gene"
/codon_start=1
/protein_id="CA04679.1"
/db_xref="GI:3879478"
/db_xref="SPTREMBL:002295"
/translation="MPHQAVDIDHPIVPPPPMDPANOAKKMSRIKESYNNK
PCQHGAPPOQAANLPHEDVLIWPARPNTNHDOSRQFGLYITAIYFLFPAAC
NSFCFNPSTLNTIMMSSESPAIIVMPYPMAPNSOAKKQSRQSRFSLKVCHEC
GALPHAQIRINLPEQVFTWPAAPRINSIQDPSQFQIVGIVIAIVLFAACVLP"
complement(join(13169..13256,13339..13401,13447..13526))
/gene="T04C12.3"
complement(join(13169..13256,13339..13401,13447..13526))
/gene="T04C12.3"
/note="predicted using GeneFinder"
/codon_start=1
/protein_id="CA04677.1"
/db_xref="GI:3879476"
/db_xref="SPTREMBL:045745"
/translation="WAVINRAVLILLCIFALSEAVSRNLEDRNMQSFQEPVGAAG
QMGGPYIHYISEYFGPRMKRHSAGSTYPSL"
complement(join(19927..20086,20140..20920,20969..21158))
/gene="T04C12.4"
complement(join(19927..20086,20140..20920,20969..21158))
/gene="T04C12.4"
/note="predicted using GeneFinder
contains similarity to Pfam domain: PF00022 (Actin),

Score=939.7, E-value=1.8e-286, N=1
 CDNA EST EMBL:M68806 comes from this gene; CDNA EST
 EMBL:M89051 comes from this gene
 CDNA EST EMBL:M214597 comes from this gene; CDNA EST
 YK55f12.3 comes from this gene
 CDNA EST EMBL:T00386 comes from this gene; CDNA EST
 EMBL:M79521 comes from this gene
 CDNA EST EMBL:T01100 comes from this gene; CDNA EST
 EMBL:T00886 comes from this gene
 CDNA EST EMBL:T02421 comes from this gene; CDNA EST
 EMBL:T01777 comes from this gene
 CDNA EST EMBL:T00437 comes from this gene; CDNA EST
 YK121h7.5 comes from this gene
 CDNA EST YK7818.5 comes from this gene; CDNA EST YK107a9.5
 comes from this gene
 CDNA EST YK55f12.5 comes from this gene; CDNA EST YK66e4.5
 comes from this gene
 CDNA EST YK72c5.5 comes from this gene; CDNA EST YK85c12.5
 comes from this gene
 CDNA EST YK89a11.5 comes from this gene; CDNA EST YK94e2.5
 comes from this gene
 CDNA EST YK65c9.5 comes from this gene; CDNA EST YK120c6.5
 comes from this gene
 CDNA EST YK70d8.5 comes from this gene; CDNA EST
 EMBL:M80032 comes from this gene
 CDNA EST EMBL:M79985 comes from this gene; CDNA EST
 EMBL:M79955 comes from this gene
 CDNA EST EMBL:T02415 comes from this gene; CDNA EST
 EMBL:M79695 comes from this gene
 CDNA EST EMBL:M79522 comes from this gene; CDNA EST
 EMBL:M79716 comes from this gene
 CDNA EST EMBL:M79981 comes from this gene; CDNA EST
 EMBL:M80068 comes from this gene
 CDNA EST CEMS874FB comes from this gene; CDNA EST
 EMBL:M79958 comes from this gene
 CDNA EST EMBL:M79631 comes from this gene; CDNA EST
 EMBL:M79632 comes from this gene
 CDNA EST EMBL:M79984 comes from this gene; CDNA EST
 EMBL:T02452 comes from this gene
 CDNA EST CEMS93FB comes from this gene; CDNA EST
 EMBL:T00475 comes from this gene
 CDNA EST EMBL:M79765 comes from this gene; CDNA EST
 YK74f5.3 comes from this gene
 CDNA EST YK85c12.3 comes from this gene; CDNA EST
 YK89a11.3 comes from this gene
 CDNA EST YK107c1.3 comes from this gene; CDNA EST
 YK108f2.3 comes from this gene
 CDNA EST YK81b4.5 comes from this gene; CDNA EST YK65c9.3
 comes from this gene
 CDNA EST YK65f1.3 comes from this gene; CDNA EST YK66e4.3
 comes from this gene
 CDNA EST YK72c5.3 comes from this gene; CDNA EST YK73a6.3
 comes from this gene
 CDNA EST YK94e2.3 comes from this gene; CDNA EST YK120c6.3
 comes from this gene
 CDNA EST YK65f1.5 comes from this gene; CDNA EST YK72h6.5
 comes from this gene
 CDNA EST YK73a6.5 comes from this gene; CDNA EST YK95d9.5
 comes from this gene
 CDNA EST YK77d1.5 comes from this gene; CDNA EST YK122c9.5
 comes from this gene
 CDNA EST YK125a11.5 comes from this gene; CDNA EST
 YK122c9.3 comes from this gene
 CDNA EST YK125a11.3 comes from this gene; CDNA EST
 YK204c8.3 comes from this gene
 CDNA EST YK369b5.5 comes from this gene
 /codon_start=1
 /protein_id="CAB04676.1"
 /db_xref="GI:3879475"
 /translation="KCDDVAVLVVNDGCKAGFAGDADPRAVPSIVGRPHRGV
 NVGGMGKDSVDEAGSKRGILTKYPIRGIYVNDDEKIMHFTFVLELRAPEEH
 PVLTLEAPLNPRANKREKMTQIMFEFNPAMYVAIOAVLSLVASGRTGVLDSDGCV

THRVPIYEGVALPHAILRLDLACRDLTDYIMKILTERGYSFTTAREIVRDIKELIC
 YVALDFEOMATATASSSSLEKSYELPDGQVITVGNFRFCPEAMPQPSFGMEACIHI
 EYVNSIMKCDIDIRKDLVYAVVLSGTTWYPCGIADMOKEITALAPSTWIKIILAP
 ERKYSWIGGSIILASTFOOMNISKOEXDESPSYHRFCF"
 complement(join(24815..24974,25027..25807,25853..26042))
 /gene="T04C12.5"
 complement(join(24815..24974,25027..25807,25853..26042))
 /gene="T04C12.5"
 /note="predicted using GeneFinder
 contains similarity to Pfam domain: PF00022 (actin),
 Score=942.3, E-value=2.9e-287, N=1
 CDNA EST EMBL:M68939 comes from this gene
 CDNA EST EMBL:T01997 comes from this gene
 CDNA EST CEMS84FB comes from this gene
 CDNA EST CEMS15FB comes from this gene
 CDNA EST EMBL:T00513 comes from this gene
 CDNA EST EMBL:M79888 comes from this gene
 CDNA EST YK77g12.3 comes from this gene
 CDNA EST YK106a8.3 comes from this gene
 CDNA EST YK107h8.3 comes from this gene
 CDNA EST YK106a8.5 comes from this gene
 CDNA EST YK107h8.5 comes from this gene
 CDNA EST YK68d3.3 comes from this gene
 CDNA EST EMBL:T02226 comes from this gene"
 /codon_start=1
 /protein_id="CAB04675.1"
 Query Match 83.3%; Score 15; DB 3; Length 35221;
 Best Local Similarity 73.3%; Pred. No. 35;
 Matches 11; Conservative 4; Mismatches 0; Indels 0; Caps 0;
 Oy 1 aaugccnauugcug 15
 ||:||||:|:||||:|
 Db 31811 AATGCCATCGGTG 31825
 RESULT 4
 AR111665/c
 LOCUS AR111665 92 bp DNA linear PAT 14-FEB-2001
 DEFINITION Sequence 3 from patent US 6127171.
 ACCESSION AR111665
 VERSION AR111665.1 GI:12828513
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 92)
 AUTHORS Sillatý,S.N. and Lebel,S.
 TITLE Cloning vector containing marker inactivation system
 JOURNAL Patent: US 6127171-A 3 03-OCT-2000;
 FEATURES
 source
 1..92
 Location/Qualifiers
 BASE COUNT 21 a 28 c 25 g 18 t
 ORIGIN
 Query Match 77.8%; Score 14; DB 6; Length 92;
 Best Local Similarity 78.6%; Pred. No. 2.3e+02;
 Matches 11; Conservative 3; Mismatches 0; Indels 0; Caps 0;
 Oy 4 ggcctatcgtggtcg 17
 ||||:|:||||:|
 Db 55 GGCCTATCGGTGCG 42
 RESULT 5
 AR111672/c
 LOCUS AR111672 249 bp DNA linear PAT 14-FEB-2001
 DEFINITION Sequence 10 from patent US 6127171.
 ACCESSION AR111672
 VERSION AR111672.1 GI:12828520
 KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 249)
TITLE Siliaty,S.N. and Lebel,S.
JOURNAL Cloning vector containing marker inactivation system
FEATURES Patent: US 6127171-A 10 03-OCT-2000;
Location/Qualifiers
1.249
source /organism="unknown"
BASE COUNT 60 a 71 c 67 g 51 t
ORIGIN

Query Match 77.8%; Score 14; DB 6; Length 249;
Best Local Similarity 78.6%; Pred. No. 2.1e+02;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 4 ggcgcuaucgugcg 17
||||:||||:||||
Db 188 GGCCTATCGGTGCG 175

RESULT 6
LOCUS AR111669 372 bp. DNA linear PAT 14-FEB-2001
DEFINITION Sequence 7 from patent US 6127171.
ACCESSION AR111669
VERSION AR111669.1 GI:12828517
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 372)
AUTHORS Siliaty,S.N. and Lebel,S.
TITLE Cloning vector containing marker inactivation system
JOURNAL Patent: US 6127171-A 7 03-OCT-2000;
FEATURES Location/Qualifiers
1.372
source /organism="unknown"
BASE COUNT 92 a 91 c 95 g 94 t
ORIGIN

Query Match 77.8%; Score 14; DB 6; Length 372;
Best Local Similarity 78.6%; Pred. No. 2.1e+02;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 4 ggcgcuaucgugcg 17
||||:||||:||||
Db 233 GGCCTATCGGTGCG 220

RESULT 7
LOCUS AR111673/c 375 bp. DNA linear PAT 14-FEB-2001
DEFINITION Sequence 11 from patent US 6127171.
ACCESSION AR111673
VERSION AR111673.1 GI:12828521
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 375)
AUTHORS Siliaty,S.N. and Lebel,S.
TITLE Cloning vector containing marker inactivation system
JOURNAL Patent: US 6127171-A 11 03-OCT-2000;
FEATURES Location/Qualifiers
1.375
source /organism="unknown"
BASE COUNT 92 a 96 c 98 g 89 t
ORIGIN

Query Match 77.8%; Score 14; DB 6; Length 375;
Best Local Similarity 78.6%; Pred. No. 2.1e+02;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 4 ggcgcuaucgugcg 17
||||:||||:||||
Db 233 GGCCTATCGGTGCG 220

RESULT 8
ID AC048161 standard; DNA; HTG; 1100 BP.
XX
AC AC048161;
XX
SV AC048161.1
XX
DT 14-APR-2000 (Rel. 63, Created)
DT 14-APR-2000 (Rel. 63, Last updated, Version 1)
XX
DE Giardia intestinalis clone K11706 strain WB-C6, LOW-PASS SEQUENCE SAMPLING.
XX
XX HTG; HTGS_PHASE0.
XX
OS Giardia intestinalis
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinales; Giardia.
XX
XX [1]
XX 1-1100
XX RA Morrison H.G., McArthur A.G., Nixon J., Eakin N.Q., Kim U., Crocker M.C.,
XX Hinkle G., Holder M.E., Sogin M.L.;
XX "Giardia: a model for ancient eukaryotic genome analysis";
XX Unpublished.
XX
XX [2]
XX 1-1100
XX RA Kim U., Morrison H.G., McArthur A.G., Nixon J., Eakin N.Q., Crocker M.C.,
XX Hinkle G., Holder M.E., Sogin M.L.;
XX Submitted (14-APR-2000) to the EMBL/GenBank/DBJ databases.
XX Josephine Bay Paul Center for Comparative Molecular Biology and Evolution,
XX Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA 02543-1015, USA
XX

CC * NOTE: This record contains 1 individual
CC * sequencing reads that have not been assembled into
CC * contigs. Runs of N are used to separate the reads
CC * and the order in which they appear is completely
CC * arbitrary. Low-pass sequence sampling is useful for
CC * identifying clones that may be gene-rich and allows
CC * overlap relationships among clones to be deduced.
CC * However, it should not be assumed that this clone
CC * will be sequenced to completion. In the event that
CC * the record is updated, the accession number will
CC * be preserved.
CC * 1 1100: contig of 1100 bp in length.
XX
XX
XX Key Location/Qualifiers
FH
FH source 1. 1100
FT /db_xref="taxon:5741"
FT /organism="Giardia intestinalis"
FT /strain="WB-C6"
FT /clone="K11706"
XX

Sequence 1100 BP: 278 A; 250 C; 329 G; 239 T; 4 other;

Query Match 77.8%; Score 14; DB 33; Length 1100;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 3 uggcgcuaucgugcg 16

```

Db      874  TGGCCTATCGGTGC 887
RESULT  9
LOCUS   XL079162
DEFINITION Xenopus laevis TGF-beta related growth factor Xnr-4 (Xnr4) mRNA,
complete cds.
ACCESSION U79162
VERSION   U79162.1
KEYWORDS African clawed frog.
SOURCE    Xenopus laevis
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 1746)
AUTHORS   Joseph E.M. and Melton D.A.
TITLE      Xnr4: a Xenopus nodal-related gene expressed in the Spemann
organizer
JOURNAL    Dev. Biol. 184 (2), 367-372 (1997)
MEDLINE    97278865
REFERENCE 2 (bases 1 to 1746)
AUTHORS   Melton D.A.
TITLE      Direct Submission
JOURNAL    Submitted (21-NOV-1996) Department of Molecular and Cellular
Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA
02138, USA

FEATURES
source      1..1746
             /organism="Xenopus laevis"
             /db_xref="taxon:8355"
             /gene="Xnr4"
             /note="TGF-beta related growth factor"
             /product="Xnr-4"
             /protein_id="AAC60127.1"
             /db_xref="GI:2072799"
             /translation="MHTYFYLILFVPGSLGINSYLIKMSKRPQDVRKRTYDS
             KDLALPLISSTYMLNLYQSFHSELNHTGEGAPSLPSNRADIIIRLAKSFDHGGSRW
             TLVDFSSLSQEEHQIAEVRDFRAFGAIAAEVAVDFLHSSCOSISGCOY
             LYGSLTGLTSSRSDDTWVTEATDIIHKWERENRKGKRYEDREKOLKIPRAKSAE
             RAYCOQNTEDQOIVMAYVSNISKERLISGTATLQDAHSKYLVMQPIQIATHRRR
             RSHIFKEHVMGKHVPADSSRTLCRAVDPEFVKQICGMSWIIHPKRYAACECE
             CSPYNESEVKNNAHYMOSLANYTKGAPEVCCVPIMSSLSMTYIDHDIAFQNH
             GMIVBECGCO"
BASE COUNT 567 a 357 c 372 g 450 t
ORIGIN

Query Match 77.8%; Score 14; DB 5; Length 1746;
Best Local Similarity 78.6%; Pred. No. 1.8e+02;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 gccuauugugucga 18
||||:||||:||||
Db 991 GGCATCGGTGCGA 1004

RESULT 10
LOCUS   AF267298
DEFINITION 2082 bp DNA linear INV 26-APR-2001
Probothriocephalus sp. KBD1 18S ribosomal RNA gene, complete
sequence.
ACCESSION AF267298
VERSION   AF267298.1
KEYWORDS Probothriocephalus sp. KBD1.
SOURCE    Probothriocephalus sp. KBD1
ORGANISM

```

```

Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
Pseudophyllidae; Diphyllobothridae; Probothriocephalus.
REFERENCE 1 (bases 1 to 2082)
AUTHORS   Kododova, I., Dolezel, D., Bruckova, M., Jirku, M., Hypsa, V., Lukes, J.
and Scholz, T.
TITLE      On the phylogenetic positions of the Caryophyllidae,
Pseudophyllidae and Proteocephallidae (Eucestoda) inferred from 18S
rRNA
JOURNAL    Int. J. Parasitol. 30 (10), 1109-1113 (2000)
MEDLINE    20453240
PUBMED     10996329
REFERENCE 2 (bases 1 to 2082)
AUTHORS   Kododova, I., Bruckova, M. and Dolezel, D.
TITLE      Direct Submission
JOURNAL    Submitted (15-MAY-2000) Institute of Parasitology, Czech Academy of
Sciences, Branisovska 31, Ceske Budejovice, CZ 370 05, Czech
Republic

FEATURES
source      1..2082
             /organism="Probothriocephalus sp. KBD1"
             /specific_host="Xenodermichthys copei"
             /db_xref="taxon:125907"
             /note="Probothriocephalus sp. ex Xenodermichthys copei
             (Pisces: Alepocephalidae) from Challenger, North Atlantic,
             collected by R.A. Bray."
             /product="18S ribosomal RNA"
BASE COUNT 476 a 464 c 614 g 525 t 3 others
ORIGIN

Query Match 77.8%; Score 14; DB 3; Length 2082;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 ugccuauuguc 16
||||:||||:||||
Db 663 TGGCCTATCGGTGC 676

RESULT 11
LOCUS   BSSPC
DEFINITION B. subtilis s10/spc operon rpmc, rpsd, rplN, rplX, rplE, rpsN genes.
VERSION   X15664.1
KEYWORDS  ribosomal protein L14; ribosomal protein L16; ribosomal protein
L24; ribosomal protein L29; ribosomal protein L5; ribosomal protein
S14; ribosomal protein S17; ribosomal protein S8; rplE gene; rplN
gene; rplX gene; rpmc gene; rpsN gene; rpsd gene; spc operon.
SOURCE    Bacillus subtilis.
ORGANISM  Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
REFERENCE 1 (bases 1 to 2437)
AUTHORS   Henkin, T.M.
TITLE      Direct Submission
JOURNAL    Submitted (23-JUN-1989) Henkin T.M., ISU Medical Centre, Dept of
Biochemistry & Molecular Biology, P O Box 33932, 1501 Kings
Highway, Shreveport LA71130, U S A
REFERENCE 2 (bases 1 to 2437)
AUTHORS   Henkin, T.M., Moon, S.H., Matheakis, L.C. and Nomura, M.
TITLE      Cloning and analysis of the spc ribosomal protein operon of
Bacillus subtilis: comparison with the spc operon of Escherichia
coli
JOURNAL    Nucleic Acids Res. 17 (18), 7469-7486 (1989)
MEDLINE    90016806
FEATURES
source      1..2437
             /organism="Bacillus subtilis"
             /strain="168"
             /db_xref="taxon:1423"
             /map="12 degrees on 360 degree map"

```

	/transl_table=11	
	/protein_id="CAA33705.1"	
	/db_xref="GI:40154"	
	/db_xref="SWISS-PROT:P12879"	
	/translation="MTDPIADMITRIRNANMYRHKL"	
BASE COUNT	805 a 457 c 540 g 635 t	
ORIGIN	:	
Query Match	77.8%; Score 14; DB 1; Length 2437;	
Best Local Similarity	78.6%; Pred. NO. 1.8e+02;	
Matches 11; Conservative	3; Mismatches 0; Indels 0; Gaps 0;	
OY	5 gccuaucgugcgca 18 : : Db 1832 GCCTATCGGTCCGA 1845	
RESULT 12		
AF122920	3839 bp DNA circular VRL 28-JUL-1999	
LOCUS	TT virus isolate US35, complete genome.	
DEFINITION	AF122920	
ACCESSION	AF122920.1 GI:5616158	
VERSION	1	
KEYWORDS	TT virus.	
SOURCE	Viruses; ssDNA viruses; unclassified ssDNA viruses.	
ORGANISM	1 (bases 1 to 3839)	
REFERENCE	Ertter,J.C., Leary,T.P., Desai,S.M., Chalmers,M.L. and Mushanwar,I.K. Analyses of TT virus full-length genomic sequences J. Gen. Virol. 80 (Pt 7), 1743-1750 (1999) 99350006 2 (bases 1 to 3839) Ertter,J.C., Leary,T.P., Desai,S.M., Chalmers,M.L. and Mushanwar,I.K. Direct Submission Submitted (25-JAN-1999) 90D, Virus Discovery, Abbott Laboratories, 1401 Sheridan Road, North Chicago, IL 60064-6269, USA Location/Qualifiers 1..3839 /organism="TT virus" /isolate="us35" /db_xref="taxon:68887" 259..708 /note="Ore2t" /product="start=1 /codon_start=1 /protein_id="AADA5649.1" /db_xref="GI:5616159" /translation="MHFSRCRKRRKRTISLPLHNSOKARBSVGRMPVPNDTIOR NWFSQFYSHTSMGCAIDFIGENHNIAAMIGRPEDONPPPGALRFLPAPPAEAEP GDRAHPWPIGGGGGNGGARGGGGDAADPADADLVAAIDAEO" 579..2882 /note="Ore1; portion of sequence previously submitted in Genbank Accession Number AF124020" /codon_start=1 /product="unknown" /protein_id="AADA5650.1" /db_xref="GI:5616160" /translation="MAYRMWRNRPRRRRRMRNRNRNRPRRRRRRRRYRRRVVN RGSGWTBRRYRWKRKKRGRKKIIICOMQPVTRCNIIVGLPLICENTVATITLCIGNTLYVLA YATSDDSYVRGPFGGWTDKFTLRILTYDEYKRFPMYWTASNEIDLGCYLTLVLYVA FRHEPVEFILTNPSPTLDIETIGSPRIHGMALNRARAIPSLKNRGKHVKIKIR VGARHFVDKWTPQDIICDMTLLTFIASADMOTPFSGPLDTIVSRQVQSMTNO LSILDNFVEHTGTODLHKRIIGHLPYNTQTQAQRVENMSATGNNNVANT NITIFDOTPEPDENGIGGPYTNSDSWKYKTGVNNKTKDIPEAKSKLYDYOTQLIG ITFGSTHRHLXCGGLSSVMLSAGRSFEFKGYTDITYNDFSDRGNNIMIDMLLI KNDSTSSTKSCLLENILMASYYGEYSKYTGDNIEHNCVCYRSBYTVOLMI DHNNIRGIVPYSLNFCNGKMKGGSLLVIPIMRAPKWPTELHKEVLFAIOAQEFYA HSDIKKYSLGKIYKFKNWGNGNPSCOVVRPCCTTOSGGRNVPRSIOAVDPYRNET ELTHAWDRFFRGEGFKAIKRNODOPIPHDTFSAGFRSBRDTEALQCSOEQCKEN	
CDS		
CDS		
TITLE	JOURNAL	
FEATURES	source	

CDS

LEPVOLKRVPMWETSOESQSEENSOKOETISQOLRDOLHOKRLMGQOLSLYQMO
RQOQONHINPILPKGLALTSISHNVI
2885..3058
/note="ORF3"
/codon_start=1
/product="unknown"
/protein_id="AA045651.1"
/db_xref="GI:5616161"
/translation="MGDPKPYNPSSNDKREYQAAKYWDREPRDLSTPEYWPAPT
PKPYNAFALNKR"

BASE COUNT 1082 a 1075 c 920 g 762 t

ORIGIN

Query Match 77.8%; Score 14; DB 14; Length 3839;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 2 auggccuacugug 15
1:11111:1111:1

Db 579 ATGGCCTATCGGTG 592

RESULT 13

RLFXGHIS 5798 bp DNA linear BCT 12-SEP-1997

LOCUS Rhizobium leguminosarum fixGHISd operon.
AJ001522

DEFINITION

VERSION AJ001522.1 GI:2398780

KEYWORDS fixg gene; fixh gene; fixi gene; fixp gene; fixs gene; promoter.
SOURCE Rhizobium leguminosarum.
ORGANISM Rhizobium leguminosarum.
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Rhizobiaceae; Rhizobium.
1 (bases 1 to 5798)
Rochepeau, P.
Direct Submission
Submitted (11-SEP-1997) Rochepeau P., Department of Biological
Sciences, University of Calgary, 2500 University Dr. N.W., Alberta,
T2N 1N4, CANADA
2 (bases 1 to 5798)
Mitsch, M.J., Rochepeau, P. and Hynes, M.F.
Characterization of the two fixGHIS operons from Rhizobium
leguminosarum strain VF39
Unpublished

JOURNAL

FEATURES

source Location/Qualifiers
1..5798
/organism="Rhizobium leguminosarum"
/strain="VF39"
/db_xref="taxon:384"
1..688
/gene="fixp"
<1..688
/gene="fixp"
/codon_start=2
/transl_table=11
/product="fixp protein"
/protein_id="CA04804.1"
/db_xref="GI:2398781"
/db_xref="SPTREMBL:O33530"
/translation="ATKGMDFSSRAELIQNLDOAKVSSQTLIRDLIAAKTVHEIDS
ALREAVAGASAKVNCAPCHGSGSGPFPINLDDMLMGDLDAIQATIAHGR
FNEIDTHASEMPFADVLDTQROYAAVYWGTLNTPSDPLAGDEGQVVDNCAC
HDDAKGAEMGAPDLADAIWIKARGEDALIRYAAAPRHGVMPAMAGRLGDTYKELT
IFVHISGGGT"
995..1008
/function="anaerobox"
1061..2629
/gene="fixg"
1061..2629
/gene="fixg"
/codon_start=1
/transl_table=11

gene

CDS

promoter

gene

CDS

/evidence=experimental
/product="fixg protein"
/protein_id="CA04805.1"
/db_xref="GI:2398782"
/db_xref="SPTREMBL:O33531"
/translation="WNLYTAPYPNIERLANEYPVNAARRNROPYAPRRKVPKRAEGR
FRRIKIMVLTGLIYLYLAPWINDDRGYPADQAILIDSSRRFFFEIIMPOEFY
VAGLIVMAGFLIVTSVAGRAMCGYCPQVWVWDLFLVYRAIAGEBNAMKLDASP
MSFAKLRKVYVTHSIWLLIGIVTGAMTFYADAPSLVLSITFGHAPAAATYTAIILT
ATYVIGIMREYOVCYCMPPRIQCAMLEDNSLYVYINDRGQRSHAKKAVNGL
SVGCVCNACVAVCPMGITIDRQOQNECTICALDCDGMKLGKRGILLAVTL
SEYSNMSLATDEGRVAVOPSRVNRNEDGYTPAIRHFMNRIIFRRVYFVAAVASG
VAMLVHLTFEREDLNVDNRNPQYVLESDESILRNGYTLRIINVPPTPRDNI SLVGL
EGATMRIPEEFKEDEARSPYVAEPDAATTLKVFVTRPVTGAINEFLFVIEDTETHADR
ATYAAAFAPGDIK"
2626..3102
/gene="fixh"
2626..3102
/gene="fixh"
/codon_start=1
/transl_table=11
/product="fixh protein"
/protein_id="CA04806.1"
/db_xref="GI:2398783"
/db_xref="SPTREMBL:O33532"
/translation="MKTSAOGFTGLHMLNSAPFGVIVANVTMAFYASSWSGIYV
PNYVAOEFMRKAAMKAMASGIEGNLSIKGHEIYDIDHDKSGSPAYDVLNKR
PVGVDHEDPILTLTKTAAGREAEHDLADGDWIVYALSRNAWSSCMRNASTIRS"
3107..5392
/gene="fixi"
3107..5392
/gene="fixi"
/codon_start=1
/transl_table=11
/product="fixi protein"
/protein_id="CA04807.1"
/db_xref="GI:2398784"
/db_xref="SWISS-PROT:O33533"
/translation="MSCCTMDAESVLNLTSTFSAEEVRILASQPLGGLGQLDLSVD
VHCGGCIATERALLTPFYKTAARVNLTAARVTCVQEEIETARTDLSKILGETNSG
YKALFTPSAESDQTRNQLLALGAVSGFAAPNMLISVSWSGADATKIDMFTMISA
MIAAPALVYAGREFPKSAMNALRHGRNMDVPLSVLSYVLSINFTVHGEHAMD
ASYLIFELFLIGRTLDHIREKARAAINGRLAPRCALLINPGSRRYAVELIAG
DEISIAGERVPVGGIVSGESDLSLIVGESPPVYVADSEYSGAMULTGSIYR
ATRIAKNSILSEITGLMEAAEGGARIRIADRATIKSYVHLALVSTLWGLFEG
DMKQAMLYAVAVLITTCPCALGLAVPVQYVAVAGELRKGIWYDGSALSTRLAETDV
AFDRTGTLTWGSSRLVAVDAMDESAALARGLAHSHRSLRVLRDTERAPISFDYV
TEIPGGIEARNGADIVRLGNAAFCGTSVPRTADSPFSEVLSKNGVDLAREFPD
TLRGACEAIDRLDAAGLETLIVSGDQTVVDMTAAHGLIDRALGSLTPROKVECOR
LNGEGRVIAWGDGINDAPALAAHVMAAPATASDIGROAADLYEFTDRDAYPEATA
VARSASLIRONFALATGYVAVPVAIAGLAPLIAAVAMSTSSITIVTNALNLF
GKRPDMIRKIGRSAEVKA"
5389..5547
/gene="fixs"
5389..5547
/gene="fixs"
/codon_start=1
/transl_table=11
/product="fixs protein"
/protein_id="CA04808.1"
/db_xref="GI:2398785"
/db_xref="SPTREMBL:O33534"
/translation="MNNLITLIPALMGIGLIALFWSLKSQYDDLOGAANRIIAE
DETDRPP"

BASE COUNT 1128 a 1740 c 1760 g 1170 t

ORIGIN

Query Match 77.8%; Score 14; DB 1; Length 5798;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 1 auggccuacugug 14

|||||:|||||:
Db 1880 AATGCCCATCGCT 1893

RESULT 14
BACRLP
LOCUS
DEFINITION
10759 bp DNA linear BCT 29-AUG-1996
Bacillus subtilis ribosomal protein (rplP)XKFFQO, rpmCB, rpsQNHMK) genes, integral membrane protein (secY) gene, adenylylate kinase (ack) gene, methionine aminopeptidase (map) gene, initiation factor 1 (ifna) gene, RNA polymerase alpha (rpoA) gene. 147971

ACCESSION
VERSION
KEYWORDS
L47971.1 GI:1044970
RNA polymerase; RNA polymerase alpha-subunit; adenylylate kinase; ack gene; ifna gene; initiation factor; initiation factor 1; integral membrane protein; map gene; methionine aminopeptidase; protein translocation; ribosomal protein; ribosomal protein B; ribosomal protein L14; ribosomal protein L15; ribosomal protein L16; ribosomal protein L17; ribosomal protein L18; ribosomal protein L24; ribosomal protein L29; ribosomal protein L30; ribosomal protein L5; ribosomal protein L6; ribosomal protein L11; ribosomal protein L13; ribosomal protein L14; ribosomal protein L17; ribosomal protein L5; ribosomal protein L8; rplE gene; rplN gene; rplO gene; rplP gene; rplQ gene; rplR gene; rplX gene; rpmC gene; rpmD gene; rpmJ gene; rpoA gene; rpsE gene; rpsH gene; rpsK gene; rpsM gene; rpsN gene; rpsQ gene; secY gene.
Bacillus subtilis.
Bacillus subtilis
Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
1 (bases 1 to 10759)
Sub.J.W., Boylan,S.A., Oh,S.H. and Price,C.W.
Genetic and transcriptional organization of the Bacillus subtilis spe-alpha region
Gene 169 (1), 17-23 (1996)
96186897

SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
FEATURES
source

Location/Qualifiers
1. 10759
/organism="Bacillus subtilis"
/strain="168 Marburg"
/db_xref="taxon:1423"
/map="12 degrees"
1..152
/gene="rplP"
<1..152
/gene="rplP"
/codon_start=3
/transl_table=11
/product="ribosomal protein L16"
/protein_id="AAB06806.1"
/db_xref="GI:1044971"
/translation="VVKPGKVLFEISGVSEVAREALRLASHKLPDKTKVKKREIIGG
ESNES"
129..342
/gene="rpmC"
129..136
/gene="rpmC"
142..342
/gene="rpmC"
/codon_start=1
/transl_table=11
/product="ribosomal protein L29"
/protein_id="AAB06807.1"
/db_xref="GI:1044972"
/translation="MKANEIRDLTAETAEIOKVKSILKEELFNLRFOLATGOLENTARIR
EYRAKAIARMKTVIRERETIAANK"
348..352
/gene="rpsQ"
348..634
/gene="rpsQ"
365..634
/gene="rpsQ"
/codon_start=1

gene
RBS
gene
CDS

/transl_table=11
/product="ribosomal protein L17"
/protein_id="AAB06808.1"
/db_xref="GI:1044973"
/translation="MSECEKRNQKRYOCHVSDKMDKTTVVETFKKHTLXGKRVKX
SKRFADHNDENQAKIGDIYKIMETPLSLTKRFRLEVEVEEAVII"
657..1043
/gene="rplN"
657..666
/gene="rplN"
675..1043
/gene="rplN"
/codon_start=1
/transl_table=11
/product="ribosomal protein L14"
/protein_id="AAB06809.1"
/db_xref="GI:1044974"
/translation="MIOGPTRLKLVADNSGARVLTIKVLGSSGSRKTANIGDIYCTVK
OATPGGVKKEVYKAVIYRTKSGARSDGSYISFDENACVILIRDDKSPRGTRIFGFP
APRLRNPMKIVSLAPEVI"
1064..1072
/gene="rplX"
1064..1392
/gene="rplX"
1081..1392
/gene="rplX"
/codon_start=1
/transl_table=11
/product="ribosomal protein L24"
/protein_id="AAB06810.1"
/db_xref="GI:1044975"
/translation="MHWKKGDQKMYISGDKGKOGTILAAFPKKDRVLVEGVNMYKH
SKPTQANPQGGISNOEAPLHVSVMPLDRTGTGVTYVTKVEDGKRVYRAKKSQVLD
K"
1402..1411
/gene="rplE"
1402..1958
/gene="rplE"
1419..1958
/gene="rplE"
/codon_start=1
/transl_table=11
/product="ribosomal protein L5"
/protein_id="AAB06811.1"
/db_xref="GI:1044976"
/translation="MNRLEKRYKKEITAPALMTKFNIDSWQVYKIEKYIVNMGVDAY
QNAKAIDSAVEELTFLAGQKPVYTRAKSLIAGFLREGMFIKAVTLRGMDYFDLQK
LISVSLPRVDFRGVSXKSPDGRGNVTLGIIKQILIEPIDYDKVTKVGMIVYTTA
NTPDEARELITGVGMPEOK"
1964..2166
/gene="rpsN"
1964..1973
/gene="rpsN"
1981..2166
/gene="rpsN"
/codon_start=1
/transl_table=11
/product="ribosomal protein L14"
/protein_id="AAB06812.1"
/db_xref="GI:1044977"
/translation="MAKKSMLAKQOQRTPKFKVOEYTRCERCGRPHSVIRKFKLCRIQF
RELAYKQIGRGVKKASW"
2181..2188
/gene="rpsH"
2181..2596
/gene="rpsH"
2198..2596
/gene="rpsH"
/codon_start=1
/transl_table=11
/product="ribosomal protein S8"
/protein_id="AAB06813.1"
/db_xref="GI:1044978"

gene
RBS
gene
CDS

	RBS	/translation="MMVMDPIADMLDIIRRNAMVRHEKLEIPASKLKEISEILIKREEF FIROVEFEDSKOGIIRFYELTYGGNNNEVITIGLARISPGRLAVAKSNEVPYRVNLGLG IALISTSGQVTLDENARKKGAEVLAYW"
	gene	2609..2619
	CDS	/gene="rplP" 2609..3165 /gene="rplP" 2626..3165 /gene="rplR" /codon_start=1 /transl_table=1 /product="ribosomal protein L6" /protein_id="AAB06814.1" /db_xref="GI:1044979"
	gene	/translation="MSRYGKKLLIEPSDVYTTLDNNTVAVKCPKGEITRFPHPMETI KVENLVLTARPSDOCKEHALHGTTRSLTGMNVGVSVSGFERGLELYGVGRASGSN KLIVNGVISHPVEIIVPEBGILEVPSQTKVVVKGTDKERVGAIAANIRAAYSPPSYKGA KGIREEGEVRRKRKGKSAAK"
	RBS	3181..3560
	CDS	/gene="rplR" 3181..3192 /gene="rplR" 3199..3560 /gene="rplR" /codon_start=1 /transl_table=1 /product="ribosomal protein L18" /protein_id="AAB06815.1" /db_xref="GI:1044980"
	gene	/translation="MITRTSKNAARLKRAHVRAKLSTGAEPPLNVERSNKHITYAOII IDVNGVTIASASTLKDPLANVESTGDISAATKVGEIYAKRAEKGISDYVFDRGGYLXK HGKYKALADAAREGLKF"
	RBS	3567..4085
	CDS	/gene="rpse" 3567..3573 /gene="rpse" 3585..4085 /gene="rpse" /codon_start=1 /transl_table=1 /product="ribosomal protein S5" /protein_id="AAB06816.1" /db_xref="GI:1044981"
	gene	/translation="MRRIDPSKLELEERLTVTNFAVAVKGGRRFPAPLAYVDGKNRG HVGGTGTAQDCEPAIRKAVEDARKNILEVWGTTTPHEITIGRFAGNILLKPSSBGG TGVIAGTGPRAVYLELAGVADIILSKLSGNPTINNMRATLIQLSELKRAEDYAKLRGSKS VEELIG"
	RBS	4084..4278
	CDS	/gene="rpmD" 4084..4090 /gene="rpmD" 4099..4278 /gene="rpmD" /codon_start=1 /transl_table=1 /product="ribosomal protein L30" /protein_id="AAB06817.1" /db_xref="GI:1044982"
	Query Match	/translation="MAKLEIFLRKRSVGIRPEDQAVYRTLGKKTNOQTVAHDNALFR GMNKVSHLVSVKQG"
	Best Local Similarity	77.8%; Score 14; DB 1; Length 10759;
	Matches	11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
OY	gccaucagcgugcga	18
DB	: :	1667
LOCUS	GCGCTATCGGTGCGA	1680
RESULT_15		
AC034257/c		
AC034257	98950 bp	DNA linear PLN 15-AUG-2000

DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
Arabidopsis thaliana chromosome I BAC F11A6 genomic sequence, complete sequence.	AC034257				
	AC034257.3	GI:8134867			
HTG.					
thale cress.					
Arabidopsis thaliana					
Eukaryotes: Viridiplantae: Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.					
REFERENCE					
1 (bases 1 to 98950)					
Federici, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Alatai, H., Nguyen, M., Lam, B., Southwick, A., Miranda, M., Brooks, S., Buehler, E., Chao, O., Chin, C., Chou, J., Choi, E., Gonzalez, A., Hong, B., Johnson-Hopson, C., Khan, S., Kim, C., Koo, T., Lee, J.M., Lenz, C., Liu, A., Liu, S., Mukharsky, N., Pham, P., Sakano, H., Schwartz, J., Shinn, P., Thaler, A., Toriumi, M., Vaysberg, M., Walker, M., Yu, G., Ecker, J., Theologis, A., and Davis, R.W.					
JOURNAL					
REFERENCE					
2 (bases 1 to 98950)					
Federici, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Alatai, H., Nguyen, M., Lam, B., Southwick, A., Bei, Q., Buehler, E., Chin, C., Chou, J., Choi, E., Dunn, P., Gonzalez, A., Hong, B., Kim, C., Koo, T., Lee, J.M., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukharsky, N., Pham, P., Sakano, H., Schwartz, J., Shinn, P., Thaler, A., Toriumi, M., Vaysberg, M., Walker, M., Yu, G., Ecker, J., Theologis, A., and Davis, R.W.					
JOURNAL					
REFERENCE					
3 (bases 1 to 98950)					
Federici, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Alatai, H., Nguyen, M., Lam, B., Southwick, A., Bei, Q., Buehler, E., Chin, C., Chou, J., Choi, E., Dunn, P., Gonzalez, A., Hong, B., Kim, C., Koo, T., Lee, J.M., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukharsky, N., Pham, P., Sakano, H., Schwartz, J., Shinn, P., Thaler, A., Toriumi, M., Vaysberg, M., Walker, M., Yu, G., Ecker, J., Theologis, A., and Davis, R.W.					
JOURNAL					
REFERENCE					
4 (bases 1 to 98950)					
Federici, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Alatai, H., Nguyen, M., Lam, B., Southwick, A., Ecker, J., Theologis, A., and Davis, R.W.					
JOURNAL					
COMMENT					
On Jun 1, 2000 this sequence version replaced gi:7922057. Bases 1-53,979 of clone F11A6 overlap with bases 73,982-127,968 of ICG BAC clone F113. (gb AC022492).					
e-mail for correspondence: arabidsequence.stanford.edu Genes with similarity to proteins in the databases are named 'putative', 'like' or 'similar to'. Genes that have EST similarity but no significant protein similarity are described as 'unknown proteins'. Genes that are annotated based only on gene prediction software are described as 'hypothetical proteins'. The gene prediction programs used to predict genes include: Grail (informatics group, Oak Ridge National Laboratory, http://compbio.ornl.gov/section/index.html), GENSCAN (Chris Burge, http://genes.mit.edu/GENSCAN/), PEXA (Victor Solovayev, http://genomic.sanger.ac.uk/gf/gf.shtml), and NetPlantGene (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, location/Qualifiers					
1. 98950					
/organism="Arabidopsis thaliana"					
/cultivar="Columbia"					
/db_xref="taxon:3702"					
/chromosome="I"					

